DOCUMENT RESUME

ED 263 588 CS 209 401

AUTHOR Beard, John D., Ed.; McNabb, Scott E., Ed.
TITLE Testing in the English Language Arts: Uses and Abuses. A Collection of Articles.
INSTITUTION Michigan Council of Teachers of English.
PUB DATE 85
NOTE 81p.
AVAILABLE FROM Michigan Council of Teachers of English, P.O. Box 892, Rochester, MI 48063 ($6.00).
PUB TYPE Collected Works - Serials (022) -- Information Analyses (070)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Bilingual Education; English (Second Language); English Instruction; Language Arts; Literacy; Reading Skills; Secondary Education; *Test Bias; Test Construction; *Testing; *Testing Problems; *Test Interpretation; *Test Reliability; Test Results; *Test Theory; Writing Evaluation; Writing Skills

ABSTRACT

Intended for teachers, this collection of articles on testing in the English language arts contains the following titles: "What Do Test Scores 'Really' Mean in Educational Policy?" by George F. Madaus; "Testing and Literacy: A Contradiction in Terms?" by Marilyn Wilson; "Throwing in the TOWL," by Mary Jane Curry; "Taking the Authority Figure Out of the Language of Testing," by Petey Young; "Implications of Parent, Teacher, and Student Perspectives on the Value of School Tests," by Sheila Fitzgerald; "Sergeant Friday Meets the Taxonomy: We Need More than Just the Facts Ma'am," by James M. Brewbaker; "Large Scale Writing Assessment: An Instructional Message," by Charles R. Chew; "Self-Assessment and the Mastery of Writing," by Edgar H. Thompson; and "Trends and Issues in the Use and Misuse of Oral Language Proficiency Testing for Bilingual Program Planning," by Josefina V. Tinajero. The appendix contains a literacy statement on testing and assessment by the Michigan Association of Speech Communication, the Michigan Council of Teachers of English, and the Michigan Reading Association. (EL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED263588

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Testing In The English Language Arts: Uses And Abuses

A Collection of Articles

Edited by: John D. Beard, Wayne State University
Scott E. McNabb, Grand Rapids Junior College

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Scott E. McNabb

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Published by: The Michigan Council Of Teachers Of English

An Affiliate Member Of The National Council Of Teachers Of English

©1985 MCTE, P.O. Box 892, Rochester, Michigan, 48063

CS 209 401

Table of Contents

Page

What do Test Scores "Really" Mean in Educational Policy? 1
By George F. Madaus

An historical study of how testing came to be so relied upon in the United States as a device of educational reform, often linked with policy statements made by boards of education. Relying on testing as a principle, an infallible arbiter of students' educational worth, can eventually corrupt the very education it is meant to improve.

Testing and Literacy: A Contradiction in Terms? 12
By Marilyn Wilson

A discrepancy exists between language arts teachers' interest in their students' literacy and what is often emphasized on tests. Life-long interest, as well as competence, in language are probably not served or encouraged by narrow, limited tests of comprehension or so-called basic skills.

Throwing in the TOWL 17
By Mary Jane Curry

The experiences of a student called "John" show how tests discriminate against some groups of children. Given a standardized test requiring original writing, John wrote much less coherently than when he composed stories on familiar topics. A discovery of test bias underscores the need to evaluate the validity of tests with respect to a student's cultural and socioeconomic environment.

Taking the Authority Figure Out of the Language of Testing 23
By Petey Young

Although society has become increasingly non-authoritarian, the language of testing contains grammatical constructions that signal readers of the presence of an authority figure. Three language changes are suggested for rewording tests to invite students to take responsibility for their own learning and be prepared to enter a non-authoritarian society.

Implications of Parent, Teacher, and Student Perspectives on the Value of School Tests 28
By Sheila Fitzgerald

Standardized tests are very incomplete measures of students' learning. Yet test scores are of increasing significance in decision making about students' placements and promotion, curriculum planning, and public policies on education. This study compares the views of parents and teachers about the values of tests and also reports the viewpoints of students on test-taking.

Page

Sergeant Friday meets the Taxonomy: We need more than just the Facts Ma'am 44
By James M. Brewbaker

While professional concern over standardized testing is commendable, greater critical attention should be directed toward teacher-made tests. Such tests too often reflect a facts-only cognitive emphasis. Better tests call for using facts to analyze, synthesize, and evaluate; such measurement bespeaks a classroom in which thinking beyond the facts matters.

Large Scale Writing Assessment: An Instructional Message 49
By Charles R. Chew

An overview of writing assessment in New York State is given, illustrating the types of writing required of students. Followed by a discussion of the evaluation process used. The criteria for evaluating writing samples across the grades point the way for instruction. The article concludes with a list of cautions for test makers.

Self-Assessment and the Mastery of Writing 55
By Edgar H. Thompson

As we try to develop better ways to assess our students' writing progress, we often forget that students need to be actively involved in their own assessment. This article describes a self-evaluation strategy that students can use to evaluate their growth as writers.

Trends and Issues in the Use and Misuse of Oral Language Proficiency Testing for Bilingual Program Planning 61

An overview of the assessment tests and procedures currently used for the identification and assessment of Limited English Proficient students is given. The test results are currently used for planning and designing the language arts program within a bilingual education context. However, there are four major problem areas resulting from the misuse of these tests.

Appendix: Literacy Statement on Testing and Assessment 74
By MASC, MCTE, MRA

Contributors 75

What Do Test Scores "Really" Mean In Educational Policy?

By
George F. Madaus

When I was first invited by the MCTE publication committee to address the topic of "uses and abuses of testing," it was like asking Jerry Falwell to speak on sin. I began to feel like a seasoned ring-master with Ringling Brothers. Maybe MCTE wished me to round up all those stalwart troopers of the testing circus -- standard score, percentile rank, grade equivalent, deviation IQ, normal curve equivalent, and that new star, logits -- and put them through their paces for one more performance under the psychometric "Big Top." After all, I am a measurement man, and a good ring-master knows his act.

However, that word **"really"** in my title kept obtruding, and I began to wonder if perhaps an honest critic of circus acts was what you wanted, rather than yet another guided tour of the "Big Top!"

More and more I realized that my technical "answers" to testing problems don't address the concerns that legislators and state board members have when they mandate tests to do all sorts of things. They are so determined to enact policy, which they feel will improve our schools, that my technical warnings about the arbitrariness of cut scores, about reliability and validity are ignored. It is the bigger political symbolic picture that really interests them. They feel that they must be seen to be doing something -- and quickly. Technical issues are but details and details are best left to civil servants. Nor do my technical answers address the questions people in the firing line, who are deeply affected by these mandated testing programs, really want answered.

Consider these questions. And certainly, while there is a technical dimension to each of them, the answers people **really** want are of a different order. The issue is one of philosophy and values, not mental measurement.

* What does a test score **really** mean to a student who after doing average work for 11 years failed the Florida Functional Literacy Test by a point or two and is told she can't graduate until she scores above the cut off?

* What do test scores **really** mean to school personnel who find out that compensatory funds, awarded last year on the basis of low student test performance, have been reduced because this year's test

scores have improved?

* What do test scores **really** mean to parents who are told their child will be pulled out of regular class and sent to a remedial class because she fell below the cut score on a state test? What do those same test scores tell the remedial teacher who receives the pupil?

* What do test scores **really** mean to a district whose certification is in jeopardy because of low test scores on a state-wide exam? What do those scores mean to the school board, administrators, teachers, children, parents and yes to the real estate agents of that town?

* What does it **really** mean to a NCAA Division I college when it finds that its star recruits combined SAT scores are 690, not 700, making them ineligible to play ball freshman year? What do those same scores mean to highly sought-after high school stars, and what, if anything, can they do about it?

The common thread linking all these questions is that the test scores are used as an administrative mechanism or device, designed to implement a particular policy. The test is the linchpin of the policy. A score on the test automatically triggers a sanction or reward which is at the heart of the policy. It's not unlike Pavlovian conditioning: the reward or sanction is meant to condition the way teachers and pupils behave.

This article focuses on the use of tests as administrative mechanisms in policy. I will first trace the evolution of testing in this country from the system wide standardized testing programs of the 40's, 50's, and early 60's, into the policy use of tests in the late 60's, 70's and 80's. Next I will discuss why policy makers turn to tests. Finally, I will discuss the effects of using tests as administrative mechanisms on the curriculum, on teaching, and on students. Here I look to history as well as to the contemporary scene for illustrations.

There are two simple ideas that I want to stress. First, when test scores are linked to important sanctions as part of policy, they can corrupt and destroy the very process they are trying to monitor and improve. Second, that using tests to implement policy at the state level shifts control over the curriculum (and hence over teaching) to the state department of education. This is a little discussed, but profound, change in the structure of American education.

One comment before I begin. Please don't interpret what I say as being "anti-testing." Quite the contrary, I believe that test results, when used in conjunction with other data about pupil performance, can help educators make educationally sound decisions about their pupils. But it is the educators using all available information including tests who must make these decisions; they should not be made mechanically, solely on the basis of a test score. Further, test results can provide important independent information to parents and the public about our schools; information they now demand and to which they are justly entitled. But once again the place of testing in policy is a matter of philosophy and values rather than mental measurement.

The Evolution of Testing Programs

In this country, unlike Europe, our system of education evolved based on local rather than state control. Circumstances of the time, the size of the

nation, and difficulties in communication and transportation helped build a uniquely democratic system of education in a network of local units. With the exception of the New York State Regents there generally were no State syllabi or State mandated examinations. These are a much more recent phenomenon.

Thus after World War I, when commercial testing--based on an individual differences model--emerged, there was a ready market for tests that could be used for local rather than state purposes. These tests were designed not to certify individuals or make inter-district comparisons, but to predict and select; to help make intra-district and inter-school comparisons; to diagnose individual learning needs; to group children; and to compare local district's performance to a national norm. By 1960, aided by the National Defense Education Act, it was the rare district that did not have a standardized, norm referenced testing program, purchasing tests and the scoring service from one of the five test publishers.

We can argue about the impact of these programs on the system, on teachers, and pupils. I admit there were abuses in the way those tests were used by some people in some districts. And for years, testing professional warned against these misuses of tests. For example, the NEA in the early 70's asked for a moratorium on such testing. I still happen to think that that was a tactical and political mistake. Nonetheless, compared to the impact of the newer state mandated testing programs like now in New Jersey, most traditional school district testing programs were relatively harmless.

In the early sixties, slowly at first but inexorably, there was a shift toward using test results in state and federal policy. At first test results were used merely to inform state and federal policy makers about the condition of education. There were a number of reasons for this initial shift. First, the concept of equality of educational opportunity evolved from a real concern about equality of inputs, resources, and access to programs into a pre-occupation with the achievement of equality of outcomes. As this pre-occupation grew, test scores began to be used as a primary indicator of educational success or failure. Second the large federal expenditures for curriculum development in math and science in the 50's and 60's eventually lead policy makers to ask for student test data as an indicator of the effectiveness of these programs. Third, advocates for minority groups began to point to the large discrepancies between the test scores of white middle class students and their constituents in order to lobby successfully for compensatory funds for programs to reduce these disparities.

Fourth, national assessment was conceived in the early 60's so that, as Ralph Tyler, its founder, put it, "Necessary information might be periodically collected to furnish a basis for public discussion and broader understanding of educational progress and problems." However, for political reasons NAEP, as it came to be called, was designed in such a way that comparisons between states were never possible. Instead NAEP developed into a broad voluntary program of testing. The present raft of reports on the state of education have relied heavily on NAEP for data about student progress.

Finally, consider the controversial first Coleman Report. The Civil Rights Act of 1964, a cornerstone of President Johnson's historic "war on poverty," required the Commissioner of Education to conduct a nationwide survey and report to the President and the Congress on the "lack of availability of equality of educational opportunity for individuals by reason of race, color, or national origin." James Coleman carried out that survey, and using a standardized test of verbal ability as his outcome measure, informed the public and their policy makers that "schools bring little influence to bear on a child's achievement that is independent of his background and general social context." Never mind

such a conclusion was counter intuitive. Coleman's test results figured prominently in the growing national debate about school effectiveness. Interestingly in the post Coleman literature on school effectiveness--which is positive and optimistic--effectiveness is viewed solely in terms of one overriding concern: how well low income students perform on multiple choice tests of basic skills. Effectiveness is reduced to this one simple criterion.

Now all of these examples -- and there are many others I have neglected -- have one thing in common. Test scores were used by policy makers for information purposes, or for lobbying. However, there was no immediate and automatic decision made about individuals, schools or districts, on the basis of the test scores, nor were rewards or sanctions linked directly to the test results.

By the early seventies the use of test scores by policy makers began to shift again. One of the first signs of this shift was the performance contracting movement. Scores on reading tests became an integral part of reimbursement formulae used to pay commercial firms which contracted with school districts to provide students with instruction in reading. Performance contracting is a throwback to "payment by results" schemes of the 19th century where teachers' salaries were tied, in part, to student test performance. About the same time policy makers in a number of states linked compensatory funding to school districts, directly to low test performance. Then, of course, we have the minimum competency movement where passing a test became a graduation or promotion requirement. In some MCT programs the tests were used as a device to identify pupils in need of remedial instruction, or as a mechanism to certify school districts.

In the past year the bloom had faded slightly on the minimum competency movement. The National Commission on Excellence in Education concludes that minimum competency examination for graduations "fall short of what is needed, and the 'minimum' tends to become the 'maximum' thus lowering educational standards for all" (20). But the Commission goes on to call for standardized tests of achievement administered at major transition points used to certify student credentials, identify the need for remediation, and identify those capable of advanced work. Again testing would play a key part in policy. Proponents of this approach, like Commissioner Bell, call this kind of testing maximum competency testing. Like dedicated anglophiles, they lust after the British "O" and "A" level secondary school certification exams. Where is Paul Revere when we really need him?

More recently still, in Florida, a state, like others, which loves to mix tests with policy, the State Merit Pay Commission is flirting with creating an educational Frankenstein by grafting merit pay to student performance. Payment by results was a monster in the 19th century and no matter how Florida might try to disguise it in the politically acceptable clothing of merit pay, genes will out. Policy makers continue to busily play leap frog but they haven't yet realized it's with a unicorn.

Instead of using tests merely to inform policy makers, the results are now being employed by them to make things happen -- automatically and mechanically -- to individuals, to schools or to districts. Tests have become an integral part of educational policy. New markets for new tests opened up, and not a few people in measurement looked around and thought they had died and gone to heaven.

Reasons Leading to Test Use in Policy

What led to this shift in the way tests were used? A perception developed

that basic skills in this country were declining; that social promotion was the rule; and that the high school diploma no longer had meaning. These perceptions were based in part on complaints from businesses, colleges and the military that disturbing numbers of high school graduates lacked both literacy and numeracy skills; that graduates could not fill out applications correctly; read directions; write letters or perform basic calculations. These complaints, along with the SAT score decline, and research evidence that seemed to indicate that our schools and compensatory programs were not effective, received wide publicity in the media. This media exposure went a long way toward undermining the public's confidence in their schools.

Further, the costs of education were rising faster than inflation while enrollments were declining. People began to question what they were getting from the schools for their local, state and federal tax dollars. Finally, there was growing nostalgia for the "good old days." Given the negative perception of contemporary schools, people began to view their own education through the hazy lens of memory, and concluded that schools were "really better then."

And so the cry went up; something must be done! But what? And why turn to tests? Ironically few seem to realize that tests as the savior, ushering in a brave new world, were in fact a medieval idea that first surfaced in 15th century Italy and later had such worthy sponsors as Adam Smith in the 18th century, and John Stuart Mill and Horace Mann in the 19th. Time doesn't permit me to outline the historical use of tests in policy but I assure you testing is nothing new under the sun and our present fascination with tests is simply another instance of forward into yesterday.

Policy makers saw the need to improve instruction. For example, in Florida the wording of legislation that mandated passing a functional literacy test in order to graduate implied that the instruction process is the key to educational improvement. The framers of that legislation gave as its purpose to "[g]uarantee to each student in the Florida system of public education that the system provides instructional programs which meet minimum performance standards compatible with the state's plan for education." However, legislators, like those in Florida and elsewhere, were shrewd enough to realize that there was little they could do directly to reform instruction. Even if it were possible to mandate a certain type of instructional program they were faced with the problem of how to monitor compliance. Testing seemed to provide the answer. Tests were seen as an available, well developed, relatively cheap and administratively simple technology that let policy makers effectively side-step the issue of dealing directly with the instructional process. By attaching important rewards or sanctions to test performance, policy makers at a gut level realized that the test was transformed into a coercive device that would influence curriculum and instruction. Unleashing the hobgoblin of diploma denial, or rentention in grade, it bullies the instructional delivery system into line to prepare students for the test. It's a classic example of winning through intimidation.

Testing had other advantages. First, it was felt that the fear of diploma denial because of a low score on a state test would motivate a target population of lazy or recalcitrant students. Second, the test could be used as a strong and useful weapon in the armament of the school when they had to tell parents that their child was not ready for promotion or graduation. The weapon metaphor used by advocates to describe this use of tests is very revealing.

Third, a test has a definite symbolic value. Policy makers had to be seen doing something overt to restore the public's confidence in their schools. Whatever their deep seated suspicions and reservations may be, the public still views testing as an objective, scientific and technological totem. To mandate a

testing program, therefore, was to be seen as directly attacking the problem of falling standards. I have had the personal experience of policy makers pointing to opinion polls as an irrefutable argument that graduation tests are quite all right simply because a majority of the public supports them. Debate over the pros and cons of testing never develops. The minds of policy makers are already made up about the value of testing. They test for testing sake and they get the public to answer yes to testing without ever having to ask a question. Notice, once more the issue is one of philosophy and values not mental measurement.

In short, testing became that amazing grace which would save the world from illiterate graduates. Testing took on missionary overtones. To parody the well-known hymn: T'was tests that made my heart to fear and tests my fear relieved. Unfortunately this theology of testing is alarmingly accurate. The test, if associated with strong accountability measures **will** bully the instructional system into line. Inevitably test scores **will** rise and policy makers **will** point to the righteousness of their gospel. If this be madness then there's method in it.

Meanwhile teacher judgment about students is cannibalized. There is here, a simple paradox. Who has the most intimate information about students' command of basic skills or mastery of subject matter? Who grades assignments day in and day out? Who sees students' blackboard work, seat work, homework, day in and day out? Who hears students recite and read day in and day out? In a word who has learned to assess students from the thousand and one signals that pass between teacher and student in the day in and day out routine of the universal classroom? Therefore, who is better able than the teacher to decide who should graduate or be promoted?

And yet, the public mistrusts the teacher's judgment, pointing to the number of students passed on from one grade to the next lacking basic skills, promoted, as they see it, solely on the basis of seat time.

Therefore, the public and its policy makers demand a surrogate, antiseptic, mechanical, decision-maker complete with excellent scientific and measurement credentials to make the decisions that they imagine teachers are unable or unwilling to make.

What has brought about this lamentable state of affairs? Are teachers unwilling to make these hard decisions? Some are, but for the vast majority of teachers, precedent taught them the bitter lesson that principals, superintendents, school boards and parents don't encourage or support them making such decisions. Rather than confronting a situation we all know to be ludicrous and directing accountability at every level, the solution seems to be to hand over crucial decision making to a surrogate teacher with a narrow range of multiple choice questions, measuring an equally narrower range of abilities, with the final mechanical decision based on a still narrow point on a numerical scale.

But I must insist again (after all I am a measurement man) that our surrogate teacher can fulfill a worthy and valuable role as a teacher's aid in these crucial decisions. But let's be clear the test is still a teacher's aid and not a Supreme Court Justice. The problem is how we convince policy makers and the public of this difference.

But the law isn't the only source of metaphor when speaking about tests. Have you noticed that the public and policy makers immediately reach for medical metaphors? The test can **diagnose** student weakness, so we can prescribe **remedial treatment**. But what doctor, in her right senses, would limit her diagnosis of the patient to one or two lab tests? Proponents might counter that's all the doctor needs if the test result was a positive biopsy. But the biopsy was ordered in the first place to confirm or deny a tentative diagnosis made on the basis of listening to the patient and making a physical exam. It is **all** of the

available information which provide the data that allows a shrewd professional instinct, gained from years of close contact with patients in the clinic, to reach a competent opinion.

So we are caught in a dilemma. By and large the public still trusts the final judgment of the doctor even though persons might seek a second opinion. But sadly the teaching profession would not seem to enjoy the same level of public confidence. Thus we have the public and its policy makers substituting a cheap, easily administered, quantifiable test score in the place of professional judgment. Let's face it, these tests used in policy are here to stay for the foreseeable future. How will teachers use the tests, or will they and their students be used by them? This is the all important question about what test scores **really** mean, not their mental measurement properties. Educators, in reacting to the use of tests in policy, must walk a shaky tightrope without the benefit of a net. Simply to attack the test is to appear self-serving, opposed both to accountability and to giving the public the independent information about the schools it demands. To silently acquiese, on the other hand, is to become an accessory before the fact. I am sorry I didn't have an easier answer for you on what the teaching profession should do. But at least let's be clear about where policy is leading us.

Tests, Policy and the Control of Education

Beyond the immediate pros and cons of using tests to allocate funds or to certify pupils or school districts there is a deeper issue: the shift in educational control and power to the State Department of Education. This issue has been inadequately addressed and the place of testing in its dynamics, poorly understood. We miss the forest for the trees.

The importance of testing in controlling the schooling process depends on two factors: first, whether an agency other than the local district controls the test; and second, whether the test has a serious impact on the individuals or the district.

History very clearly shows that where an important external exam has operated, a "tradition of past exams" quickly developed. This tradition of past exams comes to define the de-facto curriculum. Teachers see the kinds of intellectual activities which previous years' test questions required and prepare students to meet these demands. "Fine," say proponents of one philosophy. It gives teachers and pupils a clear target to shoot at; it maintains standards and so there is nothing wrong with such test preparation. "A terrible practice," say advocates of a different set of values. Teachers often spend inordinate amounts of time in simply working on test preparation, and this coaching too often is at the expense of subjects or skills not examined.

The debate, essentially one of educational philosophy, overlooks the fact that the agency which controls test content through the tradition of past exams, be they minimum or maximum competency exams, controls what is taught, how it is taught, what is learned and how it is learned.

Despite a long history of intractable resistance among American educators toward state or federal control of education, the polls show that linking tests to policies like diploma denial nevertheless enjoys widespread public acceptance (and at least in my state, I am sorry to report, widespread acceptance among teachers). However, concomitantly, the polls also indicate that the overwhelming majority of the public wants the locally elected school board to hold its traditional position of authority over education. The administrative mechanism the public wants to insure that standards are met, a state test, in

effect greatly diminishes cherished local control. You can't have your cake and eat it too. It's as simple as that. Either teachers make the painful decisions about retention and graduation, and administrators and school boards support them or a state test will do it for them at the expense of local control.

Earlier I wrote that using tests as an administrative device in policy can eventually corrupt the very process it is trying to improve. This corrupting influence can be clearly seen at work on the curriculum, on teachers and on pupils. Let's briefly look at this negative impact.

Impact on Curriculum

Reviewing the effects of external tests on the curriculum over many years and in several countries, including old mother England, it would appear that faced with a choice between objectives which are explicit in the syllabus and/or implicit in the examination questions, students and teachers invariably choose the latter.

There is evidence from a number of states that have MCT graduation programs that these tests are already narrowing the curriculum. Given our free enterprise system, publishers have begun to look at external tests in order to design materials to better train pupils to take them. Children find themselves spending more and more time filling out ditto answer sheets or work books. Deborah Meier, a successful principal of a public school in Manhattan, testified at the 1981 NIE hearing on Minimum Competency Testing that in New York City, reading instruction has come to closely resemble the practice of taking reading tests. In reading class, students using commercial materials read dozens of little paragraphs about which they then answer questions. The materials they use are more and more designed to look exactly like the tests they will take in the spring. She went on to point out that when synonyms and antonyms were dropped from the test of word meaning, teachers promptly dropped commercial material that stressed them. It is also interesting to note that nationally, sales of ditto paper are way up while sales of lined theme paper are down (Grady 1983).

When the test takes hold history tells us that the only way you can change instruction, or the curriculum, is to change the exam. It's as if Newtonian laws were operative. The gravity of the test is so much greater than that of instruction that the test changes little and instruction changes much.

Impact on Teaching

Historically, the principal negative effect on teaching is that inordinate amounts of time are often devoted to mere coaching. This can reduce teachers to functional technocrats, imparting this or that skill. Matthew Arnold commenting on the way teachers adapted to the 19th century "payment by results" scheme in Britain observed that teaching had become:

> ... a game of mechanical contrivance in which teachers will and must learn how to beat us. It is found possible by ingenious preparation, to get children through the revised code examinations in reading, writing and ciphering, without their really knowing how to read, write and cipher (Sutherland, 52).

Or consider this 1888 quote from the Superintendent of the Cincinnati Public Schools commenting on tests used in the Midwest at that time for promotion:

> In the very nature of things, the coming examination with such consequences must largely determine the character of prior teaching and study. Few teachers can resist such an influence, and in spite of it, teach according to their better knowledge and judgment. They cannot feel free, if they would .... They shut their eyes to the needs of the pupils and put their strength into what will 'count' in the examinations (White 1888, p. 518).

When I use historical illustrations like Arnold's and White's, proponents of external testing summarily dismiss them. Antique Victoriana, they point out, has little relevance today. Maybe so, but human nature, mass produced by unskilled labor, doesn't change. For example, in New Jersey the Department of Education admitted that teachers teach to the test and that the Department even provides districts with copies of previous tests that some teachers used for intensive drill work, weeks or even months before the exam (Braun 1981). Further, how do we explain the dramatic decrease in the number of pupils failing state wide competency tests? For example, in Florida the first time the test was given, 41,000 students or 36% failed; recently about 1,600 or under 2% failed. These dramatic changes have been called miracles, and touted by proponents as evidence the programs are successful. I wonder how the ghosts of Arnold and White might interpret the same Florida figures. H. G. Wells, writing in 1892, put it this way:

> The examiner pipes and the teacher must dance - and the examiner sticks to the old tune. If the educational reformers really wish the dance altered they must turn their attention from the dancers to the musicians (Wells, 382).

Impact on Students

The student is, after all, the center of the whole educational process, and the real subject of any educational program. A system of external testing is something with which the student must cope, regardless of its injustices, defects or illogicalities. Whether such coping is harmful or not depends on the system and the individual. Unfortunately, there are numerous reports of harmful effects and not nearly so many of beneficial ones. Students can be under considerable pressure and anxiety and spend considerable time cramming for and worrying about certifying exams. Those most adversely affected by certification tests tend to be students already experiencing academic problems: poor white and minority students.

As we saw earlier teachers will help most students to eventually pass their exams. But are they really better at reading, at arithmetic and at writing? I remain dubious. I am reminded of Wilfred Sheed's description in **Transatlantic Blues** of a student who was sent to a Jenkins Tutorial establishment in London which:

> Offered ... successful examination results as it might be an abortion or a forged passport... [bypassing] education altogether. Their only texts were examination papers - all the relevant ones set in the last fifty years, with odds of repetition calculated and noted as in **The Racing Form** .... Within six months, I was able to pass London matriculation without knowing any of the subjects involved; and by applying Jenkins method later, to pass every exam that ever came my way afterwards. Hence, I remain a profoundly uneducated man (117).

The Paideia Proposal recognizes that coaching by itself may get students through examinations but that they will not understand what they have memorized (Adler, 1982). Tests with important rewards or sanctions associated with them can turn our classrooms into Jenkins crammers. And, while pupils become very proficient at passing the tests, the danger is that they remain profoundly uneducated.

Conclusion

By mandating tests, policy makers create an illusion that educational quality is synonymous with performance on the tests. People are told that more and more pupils pass the tests and therefore are competent to face the demands of our society; that once again the high school diploma has meaning; that at last we have truth in labeling education. The philosophy underlying this illusion is utilitarian -- over concerned with social efficiency at the expense of the wider and deeper aspects of education. The school becomes the competency factory, the teachers the workers, the students the products, the public and the prospective employers the consumers. Passing a test becomes a stamp of approval -- a guarantee against faulty goods.

But a school is simply not a factory. Schools are not places where things are mass produced. Teachers are not assembly line workers or robots. Educated persons are not stamped out or assembled, they do not come with warranties. Passing a state mandated test is no guarantee of anything, let alone of real life success. It's a lot easier to fail a test than it is to fail in life.

The debate over the use of tests in policy is really a debate over what we want from our schools. It is a debate over educational values and competing educational philosophies, a debate over means and ends. It is not a debate on technical matters related to testing. In fact, if testing is the answer then we have done a poor job of stating the question.

So let us call things the way they are. We have taken a useful instrument and elevated it to the role of unique, infallible, arbiter of educational worth. Implicit in all that has happened is nothing less than a radical mutation in educational philosophy. We are rapidly sliding back into 19th century utilitarianism, so don't be surprised if its attendant ghosts from Dickens, Bounderby and Gradgrind, begin to stalk the land. By merely focusing on the test results we are side-stepping the more crucial question of the proper **role** of testing.

To return to my opening metaphor. No decent critic would attempt to rate a circus solely on the basis of a single performance on the tightrope while ignoring the **dexterity** of the acrobatics; the **control** of the lion-tamer; even the simple **precision** of the clowns with custard pies. Each performance has its own skills and no single area of competence exhausts the rich potential of the whole.

The human circus is equally diverse, and its assessment ranges over a broad spectrum of abilities and talents. No single activity can stand as symbolic representative of the sheer richness of human possibilities. Let us accept the wealth and diversity of educational achievement and take the time, and risks, entailed in making a full and adequate judgment upon them.

Perhaps after all, I am **not** the ringmaster and the ease with which I assumed the role at the outset is an indication of the delusions of grandeur foisted on measurement people by public expectations, and the patronage of policy makers. You teachers are the only ones who experience the whole performance and can rate its worth in the light of your experience. We can certainly devise ways and means to assist you, but ultimately the final critique must rest with you.

* * * * * * * * * * *

WORKS CITED

Adler, Mortimer J. **The Paideia Proposal, An Educational Manifesto.** N.Y.: MacMillan Publishing Co., 1982.

Braun, R. J. "Education Firm Chief warns of Stress on Skills Tests." **Newark Star Ledger**, 11 Oct. 1981, p. 80 (a).

Braun, R. J. "'Politicizing' Skills Test Scores carries Great Risk for Students." **Newark Star Ledger**, 12 Oct. 1981, p. 15 (b).

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, C., Wood, A. M., Weinfeld, F. O. and York, R. L. **Equality of Educational Opportunity.** Washington, D.C.: Office of Education, U.S. Department of Health, Education, and Welfare, 1966.

Grady, David. "The New No. 2 Pencil." **Computer Update.** (May/June) 1983:65-57.

National Commission on Excellence in Education, **A Nation at Risk: The Imperative for Educational Reform.** April, 1983.

Sutherland, G. **Elementary Education in the Nineteenth Century.** London Historical Association, 1981.

Tyler, Ralph W., "National Assessment: A History and Sociology," **School and Society** 98 (1970):471-477.

Wells, H. G. "On the True Lever of Education." **Educational Review**, 4 (November) 1892.

White, E. E. "Examinations and Promotions." **Educational Review**, 8, 1888.

Testing And Literacy: A Contradiction In Terms?

By
Marilyn Wilson

The tyranny of testing has long been acknowledged by educators. As demands for accountability increase, we teachers have gotten ourselves into a bind: teaching and testing have become a self-justifying cycle. We test what we teach, and we teach what we test. Testing demands instructional time, and even more seriously, it makes demands on the curriculum. Before long we have a curriculum dictated by the measures we choose to evaluate our students' learning.

A problem inherent in the testing of language arts is the dissonance that exists between the nature of language and literacy learning and the evaluation so often used to measure it. Language and literacy proficiencies, unlike other disciplines perhaps, are not easily defined or easily evaluated, and are not compatible with a segmental learning theory in their development or use. We learn language, reading, and writing by doing, not by being taught the bits and pieces of the process in some sequential manner. Contrary to popular opinion, reading and writing tests that segment the process into discrete skills distort the process and end up measuring something other than reading or writing proficiency.

What we evaluate and **how** we evaluate language and literacy should reflect what it is we **value**. If we value reading and writing as personal processes, we cannot evaluate them using product forms of measurement; multiple choice tests provide little insight into process. If we value errors as necessary to the developmental process, our tests must recognize their legitimacy and must not automatically penalize students for making them. And if we value reading and writing for pleasure as well as for the receiving and giving of information, our measures must not be limited to tests of information.

The continued focus on the "literary crisis" and the push for more accountability in American classrooms have made teachers of reading and writing particularly vulnerable to the demands for more standardized testing. Unfortunately the mania for quantifiable data to meet accountability demands reinforces a narrow, constricted view of literacy that pervades our classrooms, a view that is also ironically reinforced by various commissions on education. For example, **A Nation at Risk** (1983), the report of the Commission on Excellence in Education, in its call for upgrading literacy standards and academic performance, presents alarmist figures about the illiteracy rate in this country: 23 million American adults and 13 percent of all 17-year-olds are functionally illiterate.

Yet, the Commission neither provided a source for this information nor defined "functional illiteracy," a highly relative concept. Of course, focusing on numbers of Americans who are supposedly functionally **illiterate** rather than on the 207 million adults who are functionally **literate** is what makes headlines. The disputability of the focus or the figures aside, commissions' pronouncements about literacy levels despite their political nature, are likely to have widespread impact on educational policy and decision-making.

The Commission calls for higher standards to be achieved by administering more standardized tests at major transition points. It is true that quantifiable data are most expediently achieved through standardized measures, but expediency has its price, a price that the Commission is obviously unaware of. The scope of this paper is too limited to review the vast amount of literature and research on the limitations of standardized tests, except to point out that the very nature of such tests, rather than helping to uphold standards, may, in fact, be contributing to their decline.

The problem is suggested in the results of the National Assessment of Educational Progress (Petrosky, 1982) that show that high school students do not read sufficiently in depth, have difficulty with some types of inferences, and find it difficult to use the text to support their interpretations. Petrosky suggests that students' apparent inability to examine adequately the ideas they take away from their reading and to find support in the text for their interpretations may reflect current practices in testing and instruction:

> When multiple-choice testing and quick easy discussion dominate the curriculum, how can we expect anything but the most basic performance from students? When reading and writing are separated in the curriculum and when students are not encouraged to discuss or write about their reading in any extended, reasoned way, is it such a surprise that they then lack the most comprehensive thinking and analytic skills?

The one-right-answer syndrome of multiple-choice testing, besides discouraging imaginative, creative thought, may also have a negative impact on the view of reading the student has; such tests foster the image of the test or the teacher as the repository of correct information, the reader having an insignificant role in the whole process. Furthermore, such a focus discourages the development of strong self-concepts in young readers because it all but ignores the personal experience of reading -- the value of readers' personal responses and interpretations as they connect prior experience to the text to make sense of it and to give it personal relevance. And it is these personal experiences with literacy that are the very foundation of life-long reading and writing habits.

What we evaluate needs to be considered as well. Another recommendation made by the Commission on Excellence in Education is that the content of English classes in secondary schools should equip graduates to:

a) comprehend, interpret, evaluate, and use what they read;
b) write well-organized, effective papers;
c) listen effectively and discuss ideas intelligently.

On the surface one would find little to quibble about here. One could hardly disagree with the goals of reading comprehension or effective writing or effective use of language. What is disturbing, however, is not what they recommend, but what they do not.

Does the Commission assume that literacy for personal reasons has no place in academic instruction, or does it assume that although laudable, it is best dealt with in the elementary classroom, that secondary students have gone far beyond the personal to the intellectual? The NAEP results provide evidence to the contrary. Petrosky's analysis of student attitudes, habits, and practices in reading, for example, shows that 80.9 percent of the nine-year-olds said they enjoyed reading very much but only 42.4 percent of the seventeen-year-olds said so. Similarly, 53.6 percent of the nine-year-olds reported reading for enjoyment regularly, but only 32.7 of the seventeen-year-olds made the same claim. Petrosky then writes:

> While most students think it is important to read, few read for enjoyment and even fewer would choose to read a book in their spare time .... It is not so much the case that they cannot read, but that they choose not to .... Generally, they value reading most for its presentation of information, not for personal growth and pleasure (pp. 9-10.)

This primary concern for "efferent" reading, for reading for information (Rosenblatt), begins early in schools. When my first-grader brought home his first progress report, the reading objectives were the traditional ones of recognizing alphabet letters, knowing consonant and vowel sounds, mastering sight words, and reading with understanding. Ironically, in a different section of the report under Language Arts was the only objective emphasizing reading for pleasure and enjoyment, which is perhaps the most basic objective, and the foundation of all further literacy. By the time my son was in the third grade, this objective no longer appeared anywhere on the progress report. In two short years it was no longer a goal worthy of being included.

When thirty third-graders were asked to report on why they read (Dreher, 1984), their responses were largely positive and largely personal:

-- I read because it is interesting and I learn things.

-- When I'm mad and I read a happy story, I don't feel mad anymore.

-- It makes you smarter. It gives you something to do instead of walking around.

-- I read because I want to be with my friends.

-- Because I enjoy reading. I like mysteries. It also prevents me from getting bored when there is no one to play with.

Although some responses were more school-directed, e.g. "because the teacher says to," "to get to another grade," the majority indicated **self-selected** reading and **self-motivated** reasons for reading that may ultimately result in the development of life-long reading habits. But for that to happen, personal relevance needs to be encouraged and fostered in the classroom, secondary as well as elementary.

It is not that educators ignore personal purposes for reading and writing, for in elementary and middle schools there are individualized reading/writing programs, RIF programs, book clubs, pen pal organizations, etc. that encourage pleasure/recreational reading and writing. But such programs are often the "fluff," the extra activities that occur after the "skills" are acquired. What

prevents educators from using these activities as part of the basic literacy program is that their success is not easily quantifiable. If the assessment of the behavior or response is difficult, assessment will usually occur infrequently and the behavior will then be neglected during classroom teaching (Moore). A reader who enjoys and comprehends books of riddles or the **Guinness Book of World Records** but who stumbles through a basal reader will more likely have her stumbling documented than her enjoyment and comprehension of the other readings. The testing takes control, and personal reasons for literacy are ignored.

When there is little time or room in the curriculum for self-directed and self-motivated reading, when information retrieval becomes the primary goal for reading and its testing, and when testing becomes the end result of reading, readers begin to relinquish control of their own reading. A curriculum that dictates restricted and imposed purposes is one that ignores the critical connection between attitude and ability. Positive attitudes towards literacy all too frequently wane in the face of constant and constricted testing. In our haste to demonstrate our students' competence we may instead be stripping our students of the rich, personally rewarding experiences that will encourage reading well into adulthood. Personal reading is no longer personal if it is subjected to a comprehension test; both the process and the purpose are distorted.

What do we gain if we produce high school graduates who are functionally literate but who choose in most cases to avoid reading and writing activities? And how long will proficiencies in literacy be maintained if they are only rarely practiced? If the Commission used the NAEP results for conclusions about academic deficiencies, which I suspect it did, it obviously chose to ignore the report on attitudes and reading patterns and failed to see the correlation between abilities and attitudes. Until commissions on education recognize the inherent value of reading and writing for pleasure and personal growth, and until teachers are encouraged to foster those views in the classroom, we'll continue to produce unwilling and uninterested readers and writers, a far cry from the excellence the commissions say they are striving for.

How can we stop this self-defeating cycle? First, we need to reassert our views of the holistic nature of language processing, both in its acquisition and in its use. Along with this must come an insistence that students read and write -- not just in language arts classes but in all disciplines, and that they discuss with their teachers and among themselves the ideas that need sharing and exploring. We need to make clear the role of language -- oral and written -- in coming to know: we not only learn to write by writing, for example, but we learn by writing. And we need to provide students with opportunities to learn through languaging.

Second, we must reassess evaluation measures of student growth in language and literacy (Langer and Pradl). Moore also suggests an assessment of personal comprehension that occurs when readers set their own purposes for reading. Broadening the definition of comprehension from a narrow set of quantifiable reading behaviors to include personal comprehension means assessing readers' and writers' self-concepts as users of literacy, their attitudes towards literacy tasks, their purposes, and their behaviors. What this ultimately means is that we give control of the learning of reading and writing back to the students. Empowering students to make decisions about their reading and writing may be the impetus for them to develop literacy skills as naturally as possible.

Specifically, teachers can:

--make time for personal reading and writing in class, some of which

will be shared and assessed, some of which will not.

--model reading and writing for our students by reading and writing with them, and sharing with them.

--find more holistic measures for assessing literacy development, as Moore suggests, such as teacher-observation and documentation of students' responses to literacy events; the use of reading/writing conferences with individuals to assess progress and solve problems; the use of writing and reading folders by individual students as a check on topics and progress in reading and writing; and the use of questionnaires, inventories, and checklists for discovering students' interests and concerns.

Clearly what is needed is not more testing but less restrictive views of literacy and its assessment. Kids learn language in all its forms when they see purpose, when they are empowered to read and write for personal purposes, when their strengths as language users are clearly recognized and when their abilities are allowed to develop in ways compatible with natural language learning. Evaluation must be a holistic, on-going process that supports learning. Holding over students a battery of standardized tests will not ensure higher expectations or higher standards, nor will it foster the development of literacy. It will only reinforce the very weaknesses in literacy skills that educators are seeking to eliminate. Our students deserve better.

* * * * * * * * * *

WORKS CITED

Dreher, Mariam Jean. "Children's Reasons for Reading." **Reading Improvement**, 21 (1984):338-341.

Langer, Judith A. and Gordon M. Pradl. "Standardized Testing: A Call for Action." **Language Arts**, 61 (1984):764-767.

Moore, David. "A Case for Naturalistic Assessment of Reading Comprehension." **Language Arts**, 61 (1983):957-969.

A Nation At Risk: The Imperative for Educational Reform. The National Commission on Excellence in Education, April, 1983.

Petrosky, Anthony. "Reading Achievement." In A. Berger and H. Robinson (eds.). **Secondary School Reading: What Research Reveals for Classroom Practice.** National Conference on Research in English, 1982.

Reading, Thinking, and Writing: Results from the 1979-80 National Assessment of Reading and Literature (Report. No. 11-L-01). Denver, CO: National Assessment of Educational Progress, 1981. (ERIC Document Reproduction Service No. ED 209 641; 82p.)

Rosenblatt, Louise. **The Reader, The Text, The Poem.** Carbondale, Illinois: Southern Illinois University Press, 1978.

Throwing in the TOWL

By
Mary Jane Curry

John rarely smiled, but he smiled when he read his stories to my class. His narratives were the first schoolwork the fifth-grader had been enthusiastic about in the year he had attended my resource class for learning-disabled students. So I was surprised when, as I distributed a three-picture sequence of science fiction scenes to write about, he froze and at first was reluctant to tell me why he wasn't writing. Finally he said he didn't know anything about "that" (the pictures). Many of the other thirty students given this writing assignment showed similar but less extreme reactions; they had a difficult time composing a science fiction story.

John's reaction to the pictures, part of a standardized writing test, exemplifies a flaw at the base of our educational system: tests that meet criteria to be called "standardized" in most cases discriminate against some groups of children. Such is the case with the test that I gave John and his classmates in Birmingham, Alabama. It is the **TOWL (Test of Written Language**, 1978 edition, by Donald D. Hammill and Stephen C. Larsen). A series of seven writing-related tests, the TOWL includes a "Thematic Maturity" component, the only standardized test of original writing that educators of the learning-disabled consider to be valid for diagnosing writing problems. It is also used within classes of the learning-disabled to determine students' progress, which must be documented and addressed annually in Individual Education Programs (IEP's), mandated by Public Law 94-142. In the TOWL teacher's manual, the authors point out that it is a "first step" in diagnosis, and should be followed by extensive informal analysis in order to design an appropriate writing program. My argument is that this first step is, for some students, a misstep.

I discovered John's lack of familiarity with science fiction subjects when I administered the TOWL composition test as part of my research in discourse analysis. (My working hypothesis, which proved correct, was that a group of students labelled learning-disabled would write no less effectively than an equal-sized group of their nonhandicapped peers, with the two groups matched for socioeconomic level, age, and language scores on the California Achievement Test). The TOWL Thematic Maturity is an untimed test consisting of three related black-and-white drawings of outer-space fantasy scenes (see Appendix). The students silently read the following directions as the teacher reads them aloud:

> Look at the three pictures above and make up a good story to go with them. Take about five minutes to think about your story. Be sure to write a complete story using all three pictures. It is best to plan a

whole story before you begin to write. (TOWL, 1978)

The teacher scores the narrative according to twenty criteria, such as "Writes in paragraphs?" "Uses dialogue?" "Writes an integrated story using all three pictures?" "Writes a story as part of a dream sequence?" "Explains the role of the spacelings?" and "Uses some futuristic or space language?"

The small size of my study population precludes any definitive judgments about bias in the TOWL stimulus pictures, but my students' reactions to them are proof that at least one group of poor urban children, most of whom are black, do not possess the vocabularies of science fiction or real space technology. It isn't that they do not have access to television or movies. They do, but these subjects are not relevant to their lives; and as a result, they have little interest in them. They rarely watch science fiction movies or television programs, and had shown only passing interest in the Space Shuttle program.

If lack of familiarity with the subject of the pictures was an impediment for some children in the standardization studies, I could not determine it from the information in the TOWL manuals (1978 or 1983 editions). I presume not, or the authors, both widely-respected experts in learning disabilities research, would have altered the content of the pictures. I suspect instead that my students' difficulties arose because they are from a cultural and socioeconomic environment that differs significantly from that of most students participating in Hammill and Larsen's studies.

The TOWL was first tested on a population of 1,700 children in nine states, later with 3,418 in fourteen states. Proportions of students in two demographic categories, residential area (urban, suburban, or rural) and sex, parallel those reported in the U.S. Department of Commerce's 1970 and 1980 Censuses of Population. The authors collected no racial or socioeconomic data; the only southeastern states in their studies were Texas and North Carolina.

Few of my students have travelled farther than downtown Birmingham. Their parents have little or no formal education, and their cultural experiences are centered around church and their immediate neighborhoods. Some of them, despite what I view as enormous cultural deprivation, show creativity and an ability to learn as quickly as most fortunate children, given stimulation that taps their personal experiences. Such experiences are not measured by standardized English tests, which tend to emphasize isolated grammar, mechanics, and sometimes formal elements rather than the logic and semantics of a whole text.

Months of school passed before I discovered that John and his peers learned more about language through less teacher-structured assignments which allowed them to describe their world. I also found that, despite the apparent multitude of errors in their writing, I could discover more about how to teach them both traditional grammar and mechanics and the conventions of composition from their stories than I could from any formal test.

John, small compared to other fifth graders although he had failed two years, usually looked unhappy at school. He was uncommunicative with classmates as well as teachers. The principal and one of his teachers frequently reminded me that I should expect John to make little or no progress. Generally described as "spacey," "clumsy," and "stupid," he became frustrated with new lessons. But I finally hit upon something at which John could feel successful: narrative writing, sometimes accompanied by his own illustrations. Such informal writing samples proved far more useful tests of John's abilities than any formal measures I had used; more important, they gave John the means to feel competent in school. Writing what he wanted, as he wanted, John was freed from the specter of failure associated with teacher-designed tests and assignments. As his teacher I was able to assess his writing problems and strengths divorced

from any considerations of test anxiety that had apparently diminished his ability to complete acceptable work.

The first time I asked my students to write something beyond a brief letter, I was braced for John's rebelling. A group of five students was instructed to write interesting stories about any topic they chose. I would help them as much as they wanted with spelling, capitalization and punctuation. After brief individual conferences to brainstorm topics, the students -- even John -- began writing, rarely asking for help, probably because they were not familiar enough with the conventions of written language to know when they needed help. Some wrote and rewrote for an hour.

Out of this assignment, John produced the following story. I have added slash marks between sentences and given correct spellings and explanations in brackets.

> it was January 11, 1982/ we was in the house/ my father was fin ["fixin' to" -- about to] go to bed but I look up and I saw the fire/ we tride to put the fire out but we couldnot put the fire out/ a man named Sam tride to put the fire out to but he didnot put it out/ the fire department cam and they tride to put the fire out and they put the fire out/ the fireman came to my father and the fireman side [said] you tell the man [the fire chief] the story and he did/ after the fire We went to our [grandmother] Ms. Williams and we ate sun [some] ham and sun cookies and we went to bed/ the nestday [next day] it snowed and we was happy allover agin. the end.
> This is a true story.

Obviously John's story is full of mechanical and spelling errors as well as some unconventional grammatical constructions. But more significantly, it has a beginning that sets the scene, an exciting, evocative narrative in chronological order, and an ending that arrests the tension of the fire scene, the comfort John found at his grandmother's house.

Interestingly, he used the conjunctions "but" and "and" to connect more closely related units of action: "My father was fin to go to bed but I look up and I saw the fire." To shift from one unit of action, or event, to another, he simply avoided using any connective; John's substitution for end punctuation and capitalization.

In the sentence, "the fireman came to my father and said you tell the man the story and he did," John demonstrates that he can use a pronoun to refer to an antecedent unambiguously, even in a sentence with three antecedent nouns. "The man," which seemed an unclear reference to me, is often used to refer to an authority figure understood in the conversation at hand among blacks (see Frederick Erickson's discussion of black teenagers' talk in his article in **Coherence in Spoken and Written Discourse**, 1983). Over the years I have observed my black students using "the man" in conversations after the context and identity of "the man" have been established; sometimes it is used more generically, too, to refer to those in authority who exercise some power over them. John's use of the term was perfectly clear to him and his classmates. In other stories written later, except for the TOWL story, most noun and pronoun referents were readily understood.

John uses repetition of nouns more than probably recommended, but there is a strong cultural precedent, as anyone who has studied the black language in the South knows. According to Bruce Rosenberg in his study in **Folklore: Performance and Communication** (1975), repetition is a structural and cohesive device in the black sermon and in black conversations.

Except for the TOWL, John wrote several more stories that spring with logical structure and thematic coherence. He improved in mechanics and spelling and interacted more successfully with both peers and teachers. He volunteered to answer questions, even in the math class that he had previously dreaded, and he jumped at opportunities to read one of his stories. His grades on weekly tests in regular classes improved along with his behavior; other teachers willingly gave John oral tests on subjects in which he had reading difficulties.

Despite the fact that it was a test, the TOWL didn't look like a conventional test -- to me, anyway -- so I expected an impressive performance from John. The TOWL was treated as any other writing assignment: students knew by then that content, rather than spelling and mechanics, determined their grades, and they had practiced writing stories based upon illustrations, some supplied by me or another teacher, some drawn by the students themselves as preparations for writing. Contrast John's fire story above with what he wrote two months later for the TOWL:

> The men of elmers are to go seurrboom (?)/ one of the elmers man look up and see a big big fle bog Bird/ the elmers man shot at the bird/ the elmers man is mad at the bird/ the bird ent [ate?] the boy and the men the end

This attempt contains an unintelligible word or two with references inadequately explained any place in the narrative (for example, "the boy"). Events seem out of order, and there is no explanation of why the man is mad at the bird, why the bird "ent" the boy and the man. John's story barely reflects the content in the three pictures.

His raw score was 3 out of 20 possible points: he mentioned an object in the pictures; he named a planet; he gave the story a definite ending. The scaled score, based upon comparisons with those in his age group (twelve years old to twelve years and five months) was 7. His grade equivalent score (which the authors caution is to be used only for within TOWL comparisons) was 3.4. For teaching purposes, John's TOWL story was irrelevant: it was in no way indicative of his typical performance -- his strengths or his weaknesses as a writer.

The lesson I learned from this experience is that a test carrying the stamp of legitimacy isn't necessarily an appropriate measure of a particular individual's abilities. Because the contents of the TOWL stimulus pictures were somewhat unfamiliar, or at best uninteresting, to the majority of the students I tested, both the learning-disabled and nonhandicapped students' scores reflect the pictures' cultural bias rather than any written language deficits the writers might have had. Few of the thirty students scored a point for using science fiction terms, for example, or for giving names to the aliens in the stimulus pictures. None scored a point for treating the subject as part of a dream.

The TOWL Thematic Maturity test compared John's and his classmates' abilities to write a **science fiction** story to the abilities of other children across the country. My students' scores showed that they were poor science fiction writers, but those scores did not necessarily indicate that they had little imagination, or could not write dialogue, or develop characters, or organize stories into logical units. Most of them made more structural and mechanical errors while wrestling with the TOWL's unfamiliar subject matter than they made when writing about subjects of their own choosing.

As Hammill and Larsen explain in the teacher's manual accompanying TOWL, it is intended to help the teacher diagnose a student's strengths and weaknesses in

written language. The authors do suggest collecting other writing samples in order to get an accurate reading of the student's problems. However, accurate diagnosis involves more than counting T units, words per sentence, action verbs, and so on. It requires the teacher's familiarity with the child's culture and socioeconomic environment among other things. She should be familiar enough with her students to know whether a proposed test is inherently biased against them.

My argument here is that, instead of spending so much effort instructing prospective teachers on how to administer tests of limited accuracy and usefulness, colleges of education would produce more effective graduates if they made two changes: first, within testing courses, require greater emphasis upon assessing the validity of standardized tests; and second, require prospective teachers to take more courses from linguists, anthropologists, sociologists, and historians who can instill sensitivity to the factors affecting a child's modes of learning and communicating.

Teachers can glean much information from analyzing collections of students' informal compositions, yet the teacher education programs with which I am familiar include only superficial surveys of teaching written language, usually restricted to teaching grammar and mechanics rather than compositing. Requiring literature courses and the teaching of composition for special education teachers would strengthen programs for handicapped children.

Standardized tests seem entrenched in public education; within special education they exist to satisfy legal mandates to place students in special programs and to document the progress of students who have been labelled with a certain handicap. Once labelled and placed in special programs, students are often expected to perform less effectively than their nonhandicapped peers, and subsequent standardized tests they take reinforce this expectation. Students would be served better if their teachers were trained better in methods of evaluating students and in assessing the impact of these powerful instruments.

* * * * * * * * * * *

Works Cited

Erickson, Frederick. "Rhetoric, Anecdote, and Rhapsody: Coherence Strategies in a Conversation Among Black American Adolescents." **Coherence in Spoken and Written Discourse.** Ed. Deborah Tannen. Norwood, N.J.: Ablex Publishing Corporation, 1984: 81-154.

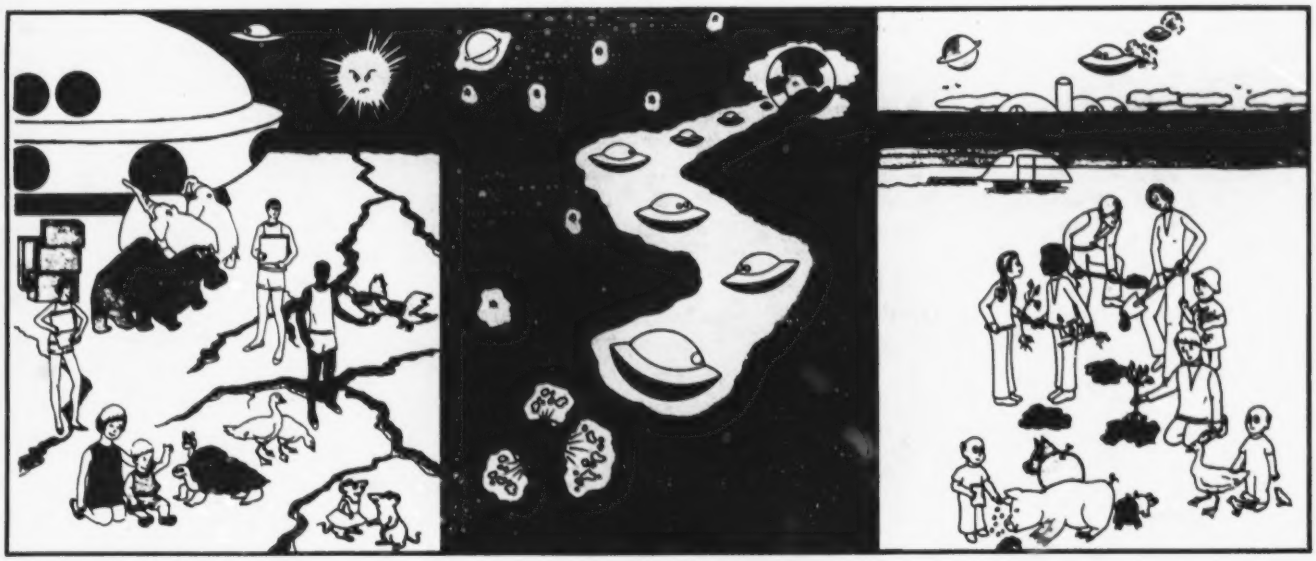
Hammill, Donald D. and Stephen C. Larsen. **Test of Written Language.** Austin, Texas: Pro-Ed, 1978, 1983.

Rosenberg, Bruce A. "Oral Sermons and Oral Narrative." **Folklore: Performance and Communication.** Ed. D. Ben-Amos and K.S. Goldstein. The Hague: Mouton, 1975: 75-101.

Appendix: TOWL TEST

Instructions: Look at the three pictures below and make up a good story to go with them. Take about five minutes to think about your story. Be sure to write a complete story using all three pictures. It is best to plan a whole story before you begin to write.

STORY



Taking The Authority Figure Out Of The Language Of Testing

By
Petey Young

> Language uses us as much as we use language. As much as our choice of forms of expression is guided by the thoughts we want to express, to the same extent the way we feel about things in the real world governs the way we express ourselves about these things.
>
> (Lakoff, 1975, p. 3)

Language symbolizes the roles people have in society and the types of performances valued by it. As the expression of a culture, language exercises power over the behavior and belief systems of those who use it (Lakoff, 1975). Although the effect of language on behavior and beliefs has been debated with some heat as recently as a decade ago, today its power is no longer questioned. For example, publishers in the field of reading refer authors to manuals outlining principles of ethnically equal and nonsexist language and refuse to publish materials containing culturally outdated language (Miller & Swift, 1980).

Language changes in education, however, tend to lag behind language changes in the surrounding culture. The lag is a necessary safety factor because education is the way people preserve culture, and consequently educators are conservative about change. Yet, in the last twenty years, America has undergone irreversible cultural changes, and now the language used in educational tests needs to begin to reflect the changes.

Cultural changes summarized in the popular book, **Megatrends: Ten New Directions Transforming Our Lives** (Naisbett, 1982), suggest that American society has gone through a cultural revolution and is no longer an industrial society but has moved beyond industrialization to become an information-processing society. The operational mode is no longer confined to the rigid structuring and authoritarian patterns necessary to an industrial society (Naisbett, 1982). In an industrial society, movement of authority occurs in well-established and relatively inflexible hierarchies from the top down. In an information-processing society, the hierarchical management structures by which the culture operates evolve into lateral networks of people with shared governance; the most valued commodity becomes information, and authority no longer flows down from the top (Deal & Kennedy, 1982).

As the culture moves in less authoritarian directions, language conventions in the society reflect a decreasing esteem for authoritarianism. For example, cooperation has become a more popular and more motivating word than competition.

Family TV serials present cooperation as a more appealing concept than competition, and, night after night, the audience warms to see human values placed over materialistic, competitive ones. Educators value process-oriented learning over product-oriented learning (Purkey & Novak, 1984). Therapists help people take authority over themselves by urging them to change from expressions that blame others to expressions in which people take responsibility for themselves. Teachers have joined this trend, and in classrooms traditional expressions such as **I have to, it makes me** are yielding to **I choose to, I am**, and **I can't** becomes **I haven't**, or **show me how** -- strong linguistic evidence that responsibility is beginning to be attributed more to individuals and less to authority figures. Yet the language of testing signals students that the world remains authoritarian by using culturally inappropriate language constructions that encourage students to feel they are not responsible for their own learning. This remains true even though researchers investigating the use and nature of classroom questioning techniques conclude that the language used by the teacher has a strong effect on the type and amount of student response (Lucking, 1985; Dillon, 1982).

The Language of Testing

The language used in testing 1) solicits product rather than process-oriented knowledge, 2) emphasizes convergent rather than divergent thinking, and 3) reminds students of the teacher's authority. Responsibility for modifying the first two belongs to individual content areas, where solutions will vary from subject to subject. However, the third is a concern across curriculum in all grades, and this paper examines three grammatical structures in which students are reminded of the teacher's authority. Such structures prepare students for an authoritarian society by carrying the message that an authority figure, who already knows the answer, is addressing the students.

The language constructions which carry authoritarian messages seem natural because they are familiar language in the American testing tradition. Fortunately, they are constructions that are linguistically easy to replace. Three changes in the language used in tests would eliminate the authoritarian voice behind the words of the test and invite students to feel responsible for their own learning. The changes involve language that keeps the teacher out and lets the student in and yet does not change the meaning of the items. All that changes is the authoritarian tone.

The Use of the Pronoun <u>You</u>

One way tests maintain an authoritarian tone is by addressing readers in the second person, indicating that an authority figure is speaking to the students. Interestingly enough, in recent years many of the questionnaires sent out by advertising corporations to consumers have been written in combinations of first and third person. For example, the directive, controlling note of questions such as, "Why did you like this soap?" have softened to "I like this soap because...," "This soap is good because...," and "I like/dislike this soap because...." These questionnaires are distributed by companies who have a considerable investment in developing instruments which consumers will fill out and return. The research behind these questionnaires has not been made public, but formats that result in a high percentage of returns would seem to merit consideration for use in the construction of tests.

The language of tests, like the language of earlier soap questionnaires, addresses readers in the second person, for example:

"Which do you consider most important?"
"Explain the reasons for your choice."
"What do you expect the result will be?"

Here the readers are being spoken to. The words contain the implication that an unseen, but well-heard, authority figure confronts the testtaker.

Pronoun reminders of the presence of the authority behind the test may be a factor in test anxiety. Certainly, what is known about the power of language (Lakoff, 1975) suggests that students' reactions to, and performances on, traditionally worded tests may be influenced by the degree to which students associate fear or anxiety with authority figures. At any rate, language reminding students that someone else has the knowledge being asked of them cannot actually be of value to the students during a test: Only the knowledge they themselves have can help them during a test. On the other hand, tests can be worded to signal students that the information is to be found inside their own heads rather than to remind them that it exists elsewhere.

One way to invite students to feel the information they seek is within their power is by omitting the pronoun **you**. Instead of wording tests in second person, tests would be worded either impersonally in the third person or, when personal reference to the reader is desired, in the first person using I, **me**, **my**, and **mine**. For example,

"When you divide 100 by 25, what quotient do you get?"
becomes
"When I divide 100 by 25, the quotient I get is..."
or
"When I divide 100 by 25, the quotient is..."
or
"When 100 is divided by 25, the quotient is...."

The Use of Commands

Commands are inherently authoritarian; they can be reworded so that students are invited to present answers rather than to respond to command. The teacher's voice disappears when commands are reworded to sound as if the ideas originate within the students' heads. For example,

"Write a paragraph explaining one cause of war."
becomes
"Here is a paragraph explaining one cause of war:...";
and
"Give the dictionary definition of audio."
becomes
"The dictionary definition of audio is...."

The Use of Question Marks

Question marks are a subtle but insidious expression of the presence of an authority. Even though most test items that end in question marks avoid the pronoun **you** and do not have a verb with an implied **you**, the nature of the questioning process itself establishes the presence of a question asker as well as a question answerer. The voice of an imaginary authority can be heard asking the question, and the insinuation is that the authority, not the student, is in control of the information. Questions are replaced by open-ended statements

which the students complete. For example,

"What is the main idea of the story?"
becomes
"The main idea of the story is..."
and
"What is the name of the capital city of Illinois?"
becomes
"The name of the capital city of Illinois is...."

Open-ended statements solicit the same information as questions but do so from inside the reader's heads without signalling the presence of outside authority.

Conclusion

Knowledge about the power of language suggests that hidden messages in the language of tests will influence the view students have of their future roles in society. The traditional language of testing contains constructions that remind students of the power of an authority figure behind the test. Such language seems designed to prepare students for life in an authority-oriented culture. However, the evidence is that American society is moving toward a less authoritarian and more laterally controlled culture (Deal & Kennedy, 1982; Naisbett, 1982). If society is moving in less hierarchical, authority-governed directions, the language of testing needs to encourage students to move in similar directions. Three easily made changes in the grammatical construction of tests would release students from linguistic reminders of authority outside themselves and invite them to assume responsibility for their own knowledge.

Models of classroom tests based upon the three changes in language have been developed at Southern Oregon State College as part of "Reading and Writing in the Secondary School," a course required for the secondary teaching certificate in Oregon. Besides the change covered in this paper, the tests also illustrate minor modifications of traditional formats, such as: 1) elimination of true/false items, 2) a set of constraints for multiple-choice items, 3) solicitation of student opinion, and 4) placement of fill-in spaces only at the ends of items where student response is less bound by syntax. The most vital component, however, is the change of tone and emphasis achieved by addressing the reader only in combinations of the neutral, third person of **it, they, its, their** and the personal, first person **I, me, my, mine.** Sample classroom tests have been developed for upper-elementary science, English, and social studies classes as well as for secondary-level general math, algebra, geometry, biology, chemistry, earth science, physics, English, history, and business -- a copy of any one is available by writing Petey Young, Department of Education, Southern Oregon State College, Ashland, OR 97520.

* * * * * * * * * * *

Works Cited

Deal, Terrence E. and Allan A. Kennedy. **Corporate Cultures: The Rites and Rituals of Corporation Life.** Reading, Massachusetts: Addison-Wesley Publishing Company, 1982.

Dillon, J.T. "Effect of Questions in Education and Other Enterprises." **Journal of Curriculum Studies, 18** (1982):64-71.

Lakoff, Robin. **Language and Society.** New York: Cambridge University Press, 1975.

Lucking, Robert. "Just Two Words." **Language Arts, 62,** (1985):173-4.

Miller, Casey and Kate Swift. **The Handbook of Nonsexist Writing.** New York: Harper & Row, 1980.

Naisbett, John. **Megatrends: Ten New Directions Transforming Our Lives.** New York: Warner Books, Inc., 1982.

Purkey, William W. and John M. Novak. **Inviting School Success: A Self-Concept Approach to Teaching and Learning.** Belmont, California: Wadsworth Publishing Company, 1984.

Implications Of Parent, Teacher, And Student Perspectives On The Values Of School Tests

By
Sheila Fitzgerald

> It would be too bad if our concentration on the implications of a decline in the statistical averages on a set of standardized examinations should seem to ignore how incomplete a measure this is of either educational or broader human purpose.
>
> **On Further Examination**
> College Entrance Examination Board
> August 1977

Background

In the 1970's and 1980's, the potential for large scale computer scoring of tests and the analysis of the test data has been realized, creating social and political forces that shape education goals and determine instructional methodologies in many states and individual school districts across the United States. In a 1977 **Phi Delta Kappa** publication, Robert Ebel stated, "The best way to determine how much learning has occurred is to observe how successfully the student can cope with tasks that require learning. This means testing." He went on to say, "If our testing and grading is educationally unproductive or harmful, it is not because the tools themselves are inappropriate. It is because we lack the skill, or are unwilling to take pains in using those skills" (8).

Typical of many editorials in newspapers in the United States in recent years is one from **The Detroit News** in February 1982 that began, "The trend toward accountability in public education is a healthy sign that school systems are finally moving in the right direction. An increasing number of districts require their high school seniors to pass a competency test prior to graduation. The practice of 'social promotion' is becoming less popular in the lower grades" (8a). In May 1981, the results of an NEA poll of 2000 teachers, a representative sample of the more than 2 million public school teachers in the United States, showed that a little more than half of the school districts provided the media with school-by-school scores on group standardized tests" (67).

The focus on tests as legitimate measures of educational outcomes is increasing, as evidenced in the reports of prestigious national commissions and committees. In **A Nation at Risk**, the report of the National Commission on

Excellence in Education, twelve of the thirteen "Indicators of Risk" were standardized test results (8-9). In its recommendations, the commission stated that, "Standardized tests of achievement should be administered at major transition points from one level of schooling to another.... The purposes of these tests are to: a) certify the student's credentials; b) identify the need for remedial intervention; and c) identify the opportunity for advanced or accelerated work" (27-28). Likewise, the Task Force on Education for Economic Growth stated that, "student progress should be measured through periodic tests of general achievement and specific skills; promotion from grade to grade should be based on mastery, not age" (11). Yet in the conclusion, the Task Force noted, "Our greatest overall educational deficiency in the United States, however, may be one that is impossible to assess through achievement tests, and impossible to measure by the usual yardsticks for gauging the adequacy of our public commitment to education. This deficiency is our absence of clear, compelling and widely agreed upon goals for improving educational performance" (31).

Many strong voices are rising to counter the increased influence of tests on schooling. "Americans," claims Carll Tucker in the **Saturday Review**, "prefer to measure everything, and what we cannot measure we prefer to believe does not exist" (56). And George Madaus, whose article leads off this collection, predicts an increase in testing for the foreseeable future, but the beginning of a backlash against the dominance of tests on policy decisions in about five years.

Many of those who criticize the influence of tests, however, also strongly question the validity and reliability of tests for what they purport to measure. Judith Langer, reporting on a two-year research project that examined the strategies students use in multiple choice reading comprehension tests, discovered that "all too frequently students selected the 'right' answer for the wrong reasons, or the 'wrong' answer for the right reasons; even more disturbing, sometimes they never had the opportunity to demonstrate their understanding of the passage at all" (765). Langer concluded that the tests do not measure the process involved in reading comprehension nor the individual student's ability to manage the comprehension process (766).

A National Academy of Sciences panel report in 1982 gave a cautionary vote of confidence to tests but warned, "Tests can only measure ability as it exists at the moment of testing; no important decision about an individual's educational future should be based on a single test score considered in isolation" (4).

Deborah Meier sees broader implications: "Standardized tests, then, are central to the current attack on the idea of democratic schooling. For all their seeming scientific sophistication and objectivity, such tests are based on a set of unexamined value judgments and social biases. As long as we rely on such tools to tell us what our children can or cannot do, we will indeed have to concede the argument to the pessimists: ordinary people cannot be taught the cognitive skills necessary for making complex decisions. In short, democracy becomes impossible" (457).

Meier raises another common criticism of test results, the veracity of the reported scores. Noting that New York City has had cheating scandals almost every year for the past five or six years, she states, "It is my belief - based on both personal knowledge and an examination of annual score fluctuations - that cheating is so rampant today that most city test scores are not even good 'norm referenced' data.... Such cheating includes everything from coaching for a specific test to actual doctoring of tests and scores" (466).

"I cannot help feeling that the current emphasis on testing and account-

ability is wrong," says Edmund Farrell of the University of Texas, "not just because it leaves to computers rather than to teachers determination of what is to be emphasized in the curriculum - we test what can be expediently scored, not necessarily what is important - but because the tests too often divert us from asking truly significant questions" (12). Stephen Tchudi, Past-President of the National Council of Teachers of English, states that the truly significant responses lie in what the student does in daily classroom work. "Actual performance rather than multiple choice test scores," he says, "should be the most important factor in judging a child's ability in school" (5).

Susan Ohanian, speaking to parents in an article in **Family Learning**, concludes, "When a parent walks into a classroom, he should look beyond these numbers (standardized test scores) to the things that really count: Is there evidence his child has done good work? Good writing? Good art? Good problem solving? Are good books being read? Is there evidence of kids' messing around, exploring on their own? Is there self-motivation, persistence, ingenuity? Is there evidence of true learning?" (5).

Design of the Study

Based on the current literature, one can predict that arguments for and against the influence of testing, discussions on appropriate uses of test data, and searches for alternatives to testing will escalate in the professional literature of education over the coming years. However, the general public -- the parents, the students and even a majority of teachers -- may not have any clear understanding of the controversies that surround the issue of testing. Indeed, many of them may, as Tucker claims, "prefer to measure everything, and believe that what cannot be measured does not exist" (52). Only a paucity of research data is available about the views of parents, teachers and pupils on tests. The eleven-item questionnaire used in this study is an attempt to draw representatives of those groups -- who are the ones most affected -- into the discussion of the role of testing in schools. This research sought their responses to the following questions:

1. What opinions do the parents of elementary, middle school and senior high school students have on the values and limitations of the standardized tests used in schools?

2. How do the opinions of teachers about the values and limitations of standardized tests differ from those of the parents?

3. Do students at three levels of schooling differ in their perspectives on the values of tests and do they experience anxiety when taking tests?

In February 1984, a survey with eleven questions about testing was sent to teacher contacts in three high schools, four middle school/junior high schools, and five elementary schools across Michigan. The schools represented city and rural settings, large and small towns, and varying socio-economic areas. Teacher contacts were asked to select three classrooms in each school, (two in one school that contained only two grades), representing lower, middle and upper grades in the school. The teachers of these classes were asked to send home with each student an eleven-item questionnaire on testing for parents to complete. In addition, the teacher was asked to select one higher ability student and one lower ability student in his/her class and conduct a five minute

open-ended oral interview with the student. (Some interviews were conducted by the teacher contact instead of the classroom teacher when the classroom teacher did not have time to conduct the interview.) Oral interviews, rather than single-response written questionnaires, were conducted to give the students open-ended opportunities for response.

In each of the schools, an eleven item questionnaire, an adaptation of the parent questionnaire, was given to each teacher in the school. Very minor changes were made on the parent questionnaire to make the items appropriate for teachers. The most significant word alteration occurred in #9:

Parents: At teacher/parent conferences, I am very interested in hearing the results of tests my child has taken.

Teachers: At teacher/parent conferences, I am careful to explain to parents the results of tests the parent's child has taken.

Three hundred seventy-eight parents, approximately 50% of the parents to whom the questionnaire was sent, returned the questionnaire (97 parents of high school students, 125 parents of middle school/ junior high school students, and 156 parents of elementary school students). It should be noted that the parent respondents may have had other children in other levels of schooling as well as at the level of the child who brought the questionnaire home; no attempt was made to have them limit their views to testing at a specific school level. One hundred eighty-five teachers completed questionnaires (86 high school teachers, 55 middle school/junior high school teachers, and 44 elementary teachers); over 50% of the teachers contacted participated in the study. Interviews were completed with 46 students (12 high school students, 16 middle school/junior high students and 18 elementary students).

Chi-square tests of significance were used to analyze the relationship between parent and teacher responses and to compare responses among the three teacher groups. Using the Statistical Package for Social Sciences, statistically significant differences were determined at p .05 and p .01 levels. The comparison of teacher responses with parent responses is summarized in Table I, and the comparison of responses for elementary, middle school, and high school teachers is given in Table II.

Implications of Tests

Table 1
Teacher-Parent Responses

Key: N=Number of responses in that category
*=p < .05
**=p < .01

Question Number	Percent of Parent Responses of True	Percent of Teacher Responses of True	x^2
	(Total = 378)	(Total = 185)	
1.	58.2% (N=216)	57.4% (N=101)	.01
2.	1 = 1.6% (N=6) 2 = 90.3% (N=237) 3 = 8.1% (N-30)	1 = 1.6% (N=3) 2 = 94.6% (N=174) 3 = 3.8% (N=7)	3.57*
3.	33.7% (N=116)	30.4% (N=48)	.41
4.	78.4% (N=291)	69.3% (N=122)	4.88*
5.	63.8% (N=224)	49.1% (N=85)	9.73**
6.	28.9% (N=106)	15.1% (N=26)	11.27**
7.	53.3% (N=176)	53.2% (N=84)	0
8.	69.4% (N=240)	40.1% (N=67)	38.88**
9.	84.2% (N=304)	63.3% (N=105)	27.55**
10.	25.1% (N=89)	18.5% (N=31)	2.51*
11.	23.3% (N=80)	17.3% (N=29)	2.08

Table 2
Teacher Responses

Key: N=Number of responses in that category
*=p < .05
**=p < .01

Question Number	Percent of High School Teachers Response of True	Percent of Middle School Teachers Response of True	Percent of Elementary School Teachers Response of True	x^2
	(Total = 86)	(Total = 55)	(Total = 44)	
1.	58.0% (N=47)	62.3% (N=33)	50.0% (N=21)	1.47
2.	1 = 3.5% (N=3) 2 = 93.0% (N=80) 3 = 3.5% (N=3)	1 = 0 % (N=0) 2 = 98.1% (N=53) 3 = 1.9% (N=1)	1 = 0 % (N=0) 2 = 93.2% (N=41) 3 = 6.8% (N=3)	5.14
3.	33.3% (N=24)	31.3% (N=15)	23.7% (N=9)	1.12
4.	74.1% (N=60)	65.4% (N=34)	65.1% (N=28)	1.6
5.	35.4% (N=28)	61.5% (N=32)	59.5% (N=25)	10.94**
6.	18.8% (N=15)	14.0% (N=7)	9.5% (N=4)	1.9
7.	51.4% (N=38)	44.4% (N=20)	66.7% (N=26)	4.33
8.	36.8% (N=28)	48.9% (N=23)	36.4% (N=16)	2.13
9.	58.3% (N=42)	67.3% (N=35)	66.7% (N=28)	1.33
10.	21.2% (N=17)	14.9% (N=7)	17.1% (N=7)	.86
11.	21.2% (N=17)	19.1% (N=9)	7.3% (N=3)	3.85

Results of the Survey

A. Parent and Teacher Questionnaires

1) The first item in the eleven item questionnaire asked parents and teachers to respond "true" or "false" to a statement about the accuracy of tests:

> I think that a test is an accurate measure of a child's ability to know and/or perform a task.

Parents and teachers were in agreement here, with nearly 60 percent of each group responding positively to item one. Only 50% of the elementary teachers agreed that tests are accurate measures of a child's ability to know and/or perform a task compared to 58% of the high school teachers and 62.3% of the middle school teachers; the differences, however, among the teacher groups was not determined to be statistically significant.

2) The second item asked how parents/teachers prefer to get information about students, in which of three ways:

> I prefer to get information about my child's learning and his/her relative standing among the children in the class from:
>
> 1. test results
> 2. a combination of test results and teacher observation
> 3. teacher observation

Most parents in this study (90.3%) prefer to get information about their children from a combination of test results and teacher observations. Only 1.6% are satisfied with test results alone. Teachers concurred, with over 94% favoring the combination, ($<.05$). Middle school parents and teachers showed the widest disparity between teacher and parent responses on this item, with 98.1% of teachers favoring the combination in contrast to 87.1% for parents, yet 10.5% of middle school parents indicated they preferred teacher observation alone and only 1.9% of teachers favored it. The differences among teachers were not statistically significant; 93% (high school), 98.1% (middle school) and 93.2% (elementary school) favored a combination of test results and teacher observations.

3) Question three asked about the frequency of tests in school:

> Tests should be given more often in school. (True or False)

A high proportion, 34 parents and 27 teachers, did not respond to this item. Only about one third (33.7%) of the responding parents thought that tests should be given more often in school, and somewhat fewer teachers agreed (30.4%). (Differences were not statistically significant.) An examination of the results of the

level of schooling shows an opposite trend between parents and teachers: Parents of high school students had the lowest percentage of parents who favored more testing (27.5%), a higher percentage given by middle school parents (34.2%) and the highest percentage given by elementary school parents (37.4%); yet the opposite trend was evident with teachers: 33.3% of high school teachers favored more testing, 31.3% of middle school teachers, and only 23.7% of elementary teachers favored more testing. The differences on this item among teachers were not statistically significant.

4) In this item, respondents were asked to indicate their judgment of tests as measures of teaching:

> Tests are not only measures of students' learning, they are measures of the effectiveness of the teaching in that classroom. (True or False)

78.4% of the parent respondents indicated that tests are measures of teaching effectiveness. Many teachers agreed, 69.3% (<.05). The results were similar for both parents and teachers, with more high school parents (84.4%) and teachers (74.1%) supporting tests as measures of effective teaching than parents of middle school students (78%) or teachers in middle schools (65.4%). Even fewer elementary parents (75%) and elementary teachers (65.1%) held this point of view. Again, the differences among the teacher groups were not statistically significant.

5) The next item asked about the importance of teaching to obtain test results:

> Teachers should teach in ways that will improve the scores on the tests that students take. (True or False)

Twenty-seven parents and 12 teachers did not respond to this item. Over two thirds of the responding parents (63.8%) thought it was important to teach to improve test results, but slightly less than one half (49.1%) of the total number of responding teachers agreed, (<.01). Both middle school parents and middle school teachers had the highest concurrence with this statement (parents 69.9%, teachers 61.5%). Only 35.4% of the high school teachers felt that they should teach to improve scores on tests. Differences among teacher groups were significant at the .01 level.

6) The publication of test results for individual schools was the subject of item six:

> It is helpful to publish test results of each school in the local newspaper. (True or False)

Less than one-third of the parents (28.9%) felt that publicizing a school's test results was helpful, and about half of that percentage of teachers (15.1%) agreed. High school parents gave the

highest approval (30.9%) and elementary teachers gave the least approval (9.5%). The differences between parents and teachers on this item were significant at the .01 level, but the teacher group differences were not statistically significant.

7) In item 7, respondents were asked if tests are more prevalent than in the past:

> I think there are more tests given to students in school now than there were 20 years ago. (True or False)

Forty-eight parents and 27 teachers did not respond to this item. Parents and teacher respondents were nearly in agreement on this issue. Slightly over half of the respondents (53.3% of the parents and 53.2% of the teachers) felt that more tests are given now than 20 years ago (no statistical difference). Elementary parents (60.2%) and elementary teachers (66.7%) were the groups who felt most strongly that testing has increased. The high school teachers (51.4%) and middle school teachers (44.4%) were less in agreement with the item, but the differences among the teacher groups were not statistically significant.

8) Item 8 asked parents and teachers if tests could be complete measures of school learning:

> There are tests for measuring most of the important things students learn in school. (True or False)

There was considerable difference between parent and teacher responses on this statement (<.01). Nearly seven out of ten parent respondents (69.4%) felt that there were tests available for most important learnings in school but only four out of ten teachers agreed (40.1%). (Non-respondents: 32 parents, 18 teachers.) Elementary parents were in least agreement among parents (64.3%), and middle school parents (73.5%) and middle school teachers (48.9%) were highest in agreement with the statement for parent and teacher groups. There was no statistically significant differences in the teacher groups.

9) The reporting of individual student's test results was the subject of this item:

> At teacher/parent conferences, I am very interested in hearing results of tests my child has taken. (True or False)

A very high percentage of parents (84.2%) want to be informed of the results of tests taken by their children. Little fluctuation exists among the three groups of parents queried. Less than two-thirds of the teachers, however, said that they were careful to report those results to parents in conferences (<.01). 67.3% of middle school teachers reported explaining test results to parents, but only 58.3% of high school teachers and 66.7% of the

elementary teachers reported doing so. Differences among the teacher groups were not statistically significant.

10) The use of test results for determining the grouping of students for instruction was the focus for item ten:

> Students should be grouped for instruction in the classroom based on test scores. (True or False)

Few parents and even fewer teachers agreed with this statement. Only one-fourth of the parents (25.1%) and less than one in five teachers (18.5%) supported tests as the basis for grouping students. The difference between parents and teachers, however, was significant at the .05 level. Middle school parents (18.9%) and middle school teachers (14.9%) showed the least agreement with the statement among the parent and teacher groups. Differences among the teachers (high school 21.2%, middle school 14.9% and elementary 17.1%) were not statistically significant.

11) This last item on the questionnaire asked respondents to make a judgement for or against tests when test results differ from a teacher's judgment:

> If a teacher's judgment on school performance is different from that of a child's performance on a standardized test, the test should hold more significance in making decisions about the child's placement in groups for instruction. (True or False)

Few parents or teachers support test results over teacher judgment when a discrepancy occurs. (34 parents and 17 teachers did not respond to the item.) Less than one-fourth of the parents (23.3%) agreed with the statement and only 17.3% of the teachers. (Difference is not statistically significant.) Elementary teachers (7.3%) were in the least agreement, with only about one in fourteen saying that tests hold significance over teacher judgment, in contrast to about one in five high school teachers (21.2%) and middle school teachers (19.1%), where the difference is not statistically significant.

B. Forty-six questionnaires were returned reporting oral interviews with students from three high schools (12 interviews), two middle school/junior high schools (16 interviews), and three elementary schools (18 interviews). The interviews elicited brief responses to five questions.

Question 1. Do you like taking tests? Why or why not?

Twenty-six of the 46 children responded positively. Some of the positive responses from each age group were the following:

High School students:
"They're challenging."

"It's interesting to see how I stand among the students."
"You know how much you learned out of the material covered."

Middle School/Junior High School students:
"It informs me what I need to know in the future."
"Tells me how much I've learned."
"It's for education so we can get jobs later, and have a good life later on."

Elementary School students:
"It's fun."
"I learn more."
"What you get graded on stays and if you do good that grade stays."

Fourteen of the 45 responses to question 1 were negative, and six of the children had mixed feelings or were undecided. Some of the negative responses for each age group included:

High School students:
"I get nervous and upset trying to do well."
"No matter how hard I study, it seems I always botch it up."

Middle School/Junior High students:
"I feel pressured, hurried, and I don't get done because I'm rushed."
"I get too nervous; I forget what I learned."

Elementary students:
"A waste of time - doesn't interest me."
"Makes me nervous."

Question 2. When you are taking a test, how do you feel? Why?

Eleven of the 46 students had positive feelings while taking a test. Some of those who responded positively said:

"Confident, because I am a good student and study for the tests."
"Relaxed; I don't get myself all upset over just a test."

Middle School/Junior High students:
"Okay, there's nothing to be afraid of."
"Good usually, because I get good grades."

Elementary students:
"I feel proud."
"Happy, because some parts I like."

Seven students expressed mixed feelings while taking tests. Twenty-eight of the 46 students expressed feelings of distress during testing sessions, such as the following:

High School students:
"Pressure - I have to think hard to remember under the pressure of time."
"The test is intimidating and I want to do well."
"I expect a lot from myself."

Implications of Tests

Middle School/Junior High students:

"I'm scared I'm going to fail it."
"I always start sweating--my hands especially."
"I feel rushed; the teacher will feel bad towards me if I get a bad grade."

Elementary students:

"Being told I have five more minutes makes me nervous."
"Afraid I might not pass."
"I don't know the percent I'm going to get."

Question 3. Why do teachers give tests?

Six of the 12 high school students indicated that the reasons teachers give tests are to evaluate what the class has learned and to base grades on results. Another three responses said that teachers want to check on whether students have been paying attention in class.

Twelve of 18 middle school students said that teachers give tests "to see how much kids have learned."

Five of 18 elementary students said that it is "to help kids learn"; four said "to see how much you have learned"; and another four students said, "to know if you are smart enough to pass."

Question 4. What do you think your teacher will do with the information from the test?

High school students had a wide range of responses to this question with little similarity. Some responses were the following:

"Depends on the teacher - some do nothing, just put it in a grade book and count it as a statistic to base grades on - others make the test a learning experience instead of just a test."
"They put results in their grade book - for some way to judge against some students."
"To evaluate his teaching."

Eight of the 16 middle school/junior high students gave responses similar to the following: "Put in their grade book - gives them information on what they can help you on." Another three students said that teachers would "use it in the future, for what to teach the kids."

Eight of the elementary students said teachers would use the test score to put it in the grade book or on the report card. Three didn't know, and two children said they would use test results "to give harder or easier work." Individual children responded, "to let parents know if their child is doing okay in school," "to give us a star" and "to publish it in the paper."

Question 5. Do you think your teacher should give you tests? Why?

All 46 of the students answered this question positively. Their explanations repeated many of the responses given for Question 3. Some other responses were:

High School students:

"I think tests are a good idea because they show both the student and the teacher how well the material is learned and how well some students

work under pressure."
"Yes, but not as constantly as they do now."
"Hopefully, they will improve our teaching and learning system."

Middle School/Junior High students:
"They are expected by the students."
"You have to be tested to see if you are doing your own work."
"If you have a test coming up you may be more apt to learn."

Elementary students:
"To see if you have been studying."
"If you don't know something, the test will find it out."
"To learn about you."

Conclusions of the Study

Parents and teachers who responded to this survey do have views on the significance of tests and how they should be used. Following are the conclusions from the responses of the adults who participated in this study:

A. Nearly six out of ten parents and teachers, especially the teachers of elementary students, believe that tests are accurate measures of learning.

B. About nine out of ten parents and teachers prefer to get information about a child's learning from a combination of test results and teacher observations rather than from tests or observations alone.

C. Only one-third of the parents believe that tests should be given more often in school. A somewhat lower percentage of teachers agree.

D. Over seven out of ten parents and nearly seven out of ten teachers agree that tests of students are measures of the effectiveness of the teaching in the classroom.

E. Over two-thirds of the parents, but slightly less than half of the teachers, believe that teachers should teach to improve test scores.

F. Fewer than three in ten parents believe that it is helpful to publish test results for each school in the local newspaper. Only about 1.5 in ten teachers find it helpful.

G. About half of the parents and the teachers, and even more parents and teachers of elementary children, think that more tests are given now than were given 20 years ago.

H. There are differences between parents and teachers about whether there are tests for measuring most of the important school learnings, with seven of ten parents believing there are, but only four of ten teachers concurring.

Implications of Tests

I. Over eight out of ten parents want to be told their child's test results, but slightly less than two-thirds of the teachers reported being careful to explain test results to parents.

J. Only one in four parents want students grouped for instruction based on test scores, and fewer than one in five teachers think this should be done.

K. Fewer than one in four parents and one in five teachers think that test results should be preferred over teacher judgments in cases of discrepancy.

A much smaller sample of students were queried by teachers for their views on testing. Although their responses were not limited to certain choices, some commonalities did emerge:

A. Over half of the students said they liked taking tests.

B. Over half of the students expressed feelings of distress while taking tests, generally pressure to perform well.

C. The most common reason teachers give tests, according to the students, is to see how much has been learned.

D. A common response for a question asking what teachers will do with test scores was "record it in a grade book or on the report card."

E. All students surveyed thought that teachers should give tests.

Implications of the Study

If, as the College Board states, tests are very incomplete measures of educational and broader human purposes, then this study shows that much more needs to be done to help parents, teachers, and students become aware of the legitimate uses of tests and, particularly, of the limitations of tests. Although it is good to see the high percentage of parents and teachers in this study who express insights into the limitations of tests, six out of ten parents and teachers believe that tests are accurate measures of learning, giving little heed to the cautions of test makers; and seven out of ten parents and teachers believe that student tests are measures of the effectiveness of teaching, a point for which there may be little support. Nearly half of the teachers and two-thirds of the parents in this study believe that teachers should teach to improve test scores, perhaps indicating a lack of awareness of the narrowness in curriculum and teaching methods that such a policy could foster. Nearly seven out of ten of the parents in the study believe that tests exist for measuring most important school learnings, yet an examination of the stated goals of most school districts would show that this is clearly not the case. Even more surprising and disturbing is the fact that 4 out of 10 teachers in this study also believe that tests exist for most important school learnings. Although 4 out of 5 teachers don't believe that students should be grouped for instruction based on test scores, the other 20% would use tests as the basis for this decision, perhaps rejecting their own insights and observations of children on a daily basis. Also, 25% of parents and 20% of the teachers in the study preferred test results over teacher judgments in cases of discrepancy, perhaps

attributing unwarranted power to paper and pencil exams given in brief time periods.

Students, as well as parents and teachers, need to receive more accurate information about the purposes and limitations of tests, as well as explanations on how their test scores will be used. Although many of the students in this study said that they had feelings of pressure during tests, this survey does not give indications that these students experienced debilitating levels of anxiety.

The results of this survey indicate that both parents and teachers value a combination of test results and teacher observations for information about children's learnings, and that parents want explanations of test results during conferences. Some teachers, however, neglect to explain tests to parents, perhaps because they do not understand the parents' interest, or they are confused about the meaning of test scores, or they may fear that test scores will counter their own professional judgments. Whatever the reason, the results of this survey indicate that some teachers need help explaining to parents the purposes and limitations of tests.

In contrast to national reports calling for more tests and an increase in statewide testing programs, two-thirds of the parents and teachers in this survey did not believe that tests should be given more often in school. If the parents and teachers in this survey are representative of parents and teachers across the country, then these groups tend to think that there are more tests given now that were given 20 years ago, especially in elementary schools, and that few tests should be given.

Publication of test results with public media is widespread. Yet the results of this survey do not support this practice: Parents and teachers strongly agree that the publishing of test results in the local newspaper is not helpful. An examination of why it is not helpful or in what ways this practice is detrimental to school programs was not a part of this study but should be investigated.

Other questions that need investigating arose from an examination of the results of this study:

- How well schooled are students who are preparing to be teachers in the purposes and limitations of tests?

- What needs to be done to help teachers understand the history, development, and implications of the growing testing industry?

- How much of the available school time is currently spent on preparation for tests, test taking, and test scoring?

- What amount of the budget is spent on the purchase of tests, the employment of testing personnel, and the dissemination of test results?

- Is the use of time and resources for testing warranted? Does the curriculum narrow to meet test expectations and ignore other important learnings?

- Why do many teachers and parents view tests as good measures of teaching effectiveness? What do parents and teachers need to know about tests?

Implications of Tests

- What should be done to help students understand tests? Do students place less significance on daily classwork than they do on tests?

The results of this study plus answers to the questions such as those listed above will help to involve students, parents and teachers in the growing debate on testing that is surfacing in the public press and in the professional journals in education.

Standardized tests offer parents, teachers and the students themselves some information about learning that occurs in school. Tests are reasonably good at measuring discrete pieces of learning, particularly whether that information can be recalled quickly in the controlled environment of the testing situation. Tests, however, measure only a small piece of the total picture of learning. Other methods of evaluation, particularly observation of the learners' application of knowledge, skills and attitudes in a variety of situations over extended periods of time, must be given priority in planning programs of instruction and measuring learning outcomes. As this survey of some parents, teachers and students indicates, the limited values of standardized tests are not completely understood. Knowledgeable educators and researchers, responsible testmakers, and questioning parents and students need to keep the issues of testing at the forefront of their discussions.

* * * * * * * * * * *

WORKS CITED

"Aptitude Tests Not Absolute." **State Journal**, Lansing, MI 14 February 1982:4.

"The Basics of Testing." **Today's Education** (Feb./Mar.) 1982:67.

Ebel, Robert L. **The Uses of Standardized Testing.** Bloomington, Indiana: Phi Delta Kappa Educational Foundation, 1977.

Ervin, Chris. "Prof Advocates More Teaching, Less Testing." **M.S.U. News Bulletin**, East Lansing, MI 18 Feb. 1982:5.

Farrell, Edmund. "The Vice/Vise of Standardized Testing: National Depreciation by Quantification." **English Journal** 65 (October 1976):10-12.

Langer, Judith. "Standardized Testing: A Call for Action." **Language Arts** 61 (Oct. 1984):764-6.

Meier, Deborah. "Why Reading Tests Don't Test Reading." **Dissent** (Fall 1981): 457-66.

National Commission on Excellence in Education. **A Nation at Risk.** Washington: U.S. Department of Education. 1983.

Ohanian, Susan. "Achievement Test Scores, Facts that Distort the Truth." **Family Learning** (Mar/Apr, 1984):52-54.

"Our Opinions: Testing Teachers." **Detroit Free Press** 15 Feb. 1982:8A.

Taskforce on Education for Economic Growth. **Action for Excellence.** Washington: Education Commission of the States, 1983.

Tucker, Carll. "The Data Game." **Saturday Review** 4 Feb. 1978:56.

* * * * * * * * * * *

SELECTED BIBLIOGRAPHY

Anderson, Beverly and Chris Pipko. "State-Mandated Testing and the Fate of Local Control." **Phi Delta Kappan** 66 (Nov. 1984):209-12.

Brady, Elizabeth H. "To Test or Not to Test." **American Educator** 1, (Winter 1977):3-9.

Fitzgerald, Sheila M. "What Are the Effects of Tests?" **Childhood Education**, 56 (Feb./Mar. 1980):216-17.

Herndon, Enid B. **Your Child and Testing.** Washington: National Institute of Education, 1981.

Owen, David. **None of the Above.** Boston: Houghton Mifflin, 1985.

Perrone, Vito. **The Abuses of Standardized Testing.** Bloomington, Indiana: The Phi Delta Kappa Educational Foundation, 1977.

Perrone, Vito, ed. **Testing and Evaluation: New Views.** Washington: Association for Childhood Education International, 1975.

Shulman, Lee S. **Test Design: A View from Practice.** East Lansing, MI: MSU Institute for Research on Teaching, 1978.

Strenio, Andrew A. Jr. **The Testing Trap.** New York: Rawson Wade, 1981.

Testing. Washington: National Education Association, 1973. ERIC ED 087 739.

Venezky, Richard L. **Testing in Reading: Assessment and Instructional Decision Making.** Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills, 1974.

Sergeant Friday Meets The Taxonomy: We Need More Than Just The Facts Ma'am

By
James M. Brewbaker

> The load of data purveyed in a typical high school year is staggering. Just because the facts roll out, of course, does not mean they are either understood or retained. Most are ignored. A few may be deliberately identified as vital for the various tests, but many, if not most, of these facts mean little to the student.
>
> (Theodore Sizer, **Horace's Compromise**)

> ... while students learn to read a wide range of material, they develop very few skills for examining the nature of the ideas that they take away from their reading. Though most have learned to make simple inferences about such things as a character's behavior and motivation, for example, and could express their own judgments of a work as 'good' or 'bad,' they generally did not return to the passage to explain the interpretations they made.
>
> (National Assessment of Educational Progress, **Reading Thinking and Writing**)

Theodore Sizer and NAEP are two among numerous authorities that have brought to light an apparent problem with the intellectual climate in secondary schools today. They point to a fact-bloated curriculum where students don't and perhaps can't do much with the facts they acquire, however briefly. Somewhat less apparent but equally telling a problem is the way teachers limit the intellectual demands of their classroom testing.

Before exploring this problem, spend a few minutes taking the quiz which follows. Imagine six classrooms, where six teachers of sophomore English, Level 2 have been working with simile and metaphor. Their honest hope is that students will grasp these elemental uses of figurative language well enough to see how they work in both literature and everyday discourse. Now it is time to find out who learned what and how well. But these six teachers of sophomore English, Level 2, despite common curriculum and text materials, elect to measure what happened to their students along the instructional way in six strikingly differ-

ent ways. We can infer much about their teaching from their approaches to evaluation.

Read each measurement item below carefully. Then, in the space provided, assign each a taxonomic level as either 1.0, Knowledge; 2.0, Comprehension; 3.0, Application; 4.0, Analysis; 5.0, Synthesis; or 6.0, Evaluation. There are no repeats, with one measurement item per taxonomic category.

Taxonomic Level	
_____	Teacher A, Mr. Arthur: Assume that a friend does not understand the definition of metaphor on p. 62 of your book. In the space provided, explain what metaphor is in words you and your friend would both understand.
_____	Teacher B, Mrs. Barron: Attached are two examples of writing with several similes in each. Which, in your view, makes the better use of this figure of speech? Why do you think so? **(One sample uses cliched similes; the other does not.)**
_____	Teacher C, Mr. Cinch: Define **metaphor** and give two examples we discussed in class.
_____	Teacher D, Miss Debono: In the space provided, write one **simile** comparing Ronald Reagan to something else and one **metaphor** comparing the school football team to something else.
_____	Teacher E, Ms. Edwards: Attached is a xerox copy of page 1 of Sunday's sports section. Using headlines only, find five similes or metaphors. For each, indicate (a) the element being compared, and (b) the element to which it is being compared.
_____	Teacher F, Mrs. Fox: Invent a new product name for a **soap**, a **deodorant**, or **shoes**. Make the product's brand name a metaphor. Then write a commercial which is an **extended metaphor** favorably comparing the product to something else.

Ready to check your answers? Mr. Arthur has constructed a **comprehension** item for his students. By translating book talk into student talk ("words you and your friend would both understand"), students successful on this task will demonstrate knowledge that has a personal dimension. Mrs. Barron has devised a sophisticated measurement task calling for cognitive **evaluation**. Students in her class apparently have also studied cliched use of language, have learned that fresh comparisons are normally preferred, and now will try to show their ability to distinguish between the two.

In contrast, Mr. Cinch plans to measure student learning at the most basic level, that of **knowledge**. Even in the examples, his sophomores will cite those "we discussed in class." Miss Debono is concerned with **application**, reasoning that a student's success in framing a simple simile and a metaphor indicates understanding of each concept.

Ms. Edwards has formulated an exercise which sends students on a metaphor hunt calling for **analysis** of newspaper headlines. Unfamiliar with the examples

in the local sports page, her students will ferret out the poetic nuances of "Lions Devour Saints" and "Vikings Pillage Raider's Camp," thus demonstrating their feel for figurative language in popular media.

Mrs. Fox, finally, has devised an activity at the **synthesis** level through which, indirectly at least, she will find out how well students can use what they know of metaphor in creating an original commercial. Her students won't necessarily know they're being tested, an arrangement which has both advantages and disadvantages.

Which teachers, would you say, are superior measurers? Your answer depends on the relative weight you give such elements as objectivity, efficiency, the student's prior learning and testing experiences, and instructional purpose.

Certainly Mr. Cinch (knowledge) and, to a lesser extent, Mr. Arthur (comprehension) and Miss Debono (application) can size things up pretty quickly, put a few numbers in their grade books, and move on. Mr. Cinch will feel good, perhaps, about his objectivity, despite the view of many professionals that such test items, precisely framed though narrowly conceived, are as basically subjective as essay items. Miss Debono (application) should be able to check off rights and wrongs with reasonable accuracy and speed.

The other three teachers, by choosing classroom measurement which many would regard as subjective, may open themselves to criticism from those who value exactness more than a test situation in which learners use facts to complete tasks at a higher cognitive level. If nothing else, the marking and grading process may become sticky. Miss Barron (evaluation) may find that what is cliched to her seems original and clever to several of her students; too, students may select the right passage but offer weak explanations as to why. Without a rubber yardstick, or because of one, she may have trouble justifying her grades even when she is convinced of their soundness. If Ms. Edwards' (analysis) students have worked with headlines before, especially headlines as metaphor, they may accomplish the task she has set; if not, she'll spend as much time explaining what to do as the kids will in completing their work. And Mrs. Fox's (synthesis) students . . . well, there's no telling what direction they'll take. If she assigns specific criteria for judging their commercials (three elaborations of the basic comparison, for example), she could gain some standardization of their creations but would probably sacrifice spontaneity at the same time. She also may have trouble coming up with letter or numerical grades she is comfortable with, and she should be ready to defend her product-oriented measurement both to querulous students and her principal.

But there is a paradox in all this. Despite the foregoing testing "problems," the Barrons (evaluation), Edwards (analysis), and Foxes (synthesis) are providing superior measurement. Many times good measurement is messy, involving as it does professional judgment of complex, varied behaviors. Much to their credit, the testing of Barron, Edwards, and Fox, for one thing, bespeaks a clearly intellectual intent and classroom environment. Facts, to these teachers, appear to be the raw material of learning -- elements to use in other activities -- rather than ends in and of themselves. What is more, through their testing, these teachers engage students in processes valuable in their own right.

A generation ago, the romantic critics claimed that children, ultimately, learn what we test. A more accurate way of putting this idea is to say that they learn **how** we test, and, when our tests are consistently and needlessly low level, our students respond in kind. In a sense, therefore, our classes are "about" both what and how we test, and the tests we use mirror our true objectives more accurately than either the system's curriculum guide or our individual lesson plans. The point is this: that learning simile and metaphor

means one thing in Mr. Cinch's class (regurgitation) and something else, something better, in Mrs. Fox's (creating new discourse built from classroom learnings).

We have been rightly concerned about the excesses of standardized tests and state-mandated competency requirements. But one problem of testing is much closer to home, in many of our own classrooms. Systems models of teaching and narrowly applied behaviorism -- each frequently espoused by ed psych professors, reinforced by teacher's guides and curriculum materials, and followed by many of our colleagues -- promote multiple-choice, lower level testing for its reliability while overlooking its impact on instruction. Essay tests and direct observation of behavior, the prevailing argument remains in too many circles, are just plain sloppy. Yet it is these so-called sloppy techniques which promote thinking beyond the facts for many of our students.

Too many teachers -- most, according to Theodore Sizer's study for the National Association of Independent Schools -- have ended up teaching and testing pretty much like Mr. Cinch. Facts, facts, and more facts, learned for the moment, learned at the most shallow level, learned one year and then learned again the next. Just the facts, ma'am.

But the times demand more than this. Facts without connections aren't enough. The intellectual stew in too many classes is bland. Sure, some of the kids may end up as whizzes at "Trivial Pursuit," that symptomatic American game, but too few feel moved, much less are able, to construct a well reasoned letter-to-the-editor. By dependence on tests which rely too heavily on level -1.0 "knowing" and too little on level 5.0 "creating," too many teachers allow themselves to become part of the problem.

In the eighties, educational critics and distinguished panels are discovering thinking skills as a new curriculum basic, and, with the discovery, a new crisis emerges. NAEP data reveal that thinking isn't what it used to be among our seventeen-year-olds. "Students ... seemed genuinely puzzled at requests to explain or defend their points of view," reports NAEP in **Reading, Thinking and Writing** (3). Sizer takes on coverage of content as the most damaging curriculum bugaboo of our time, arguing that thoughtful engagement with knowledge matters far more, say, than covering the anthology from Captain John Smith to Alice Walker, that putting facts together takes time and thought, and that educational leaders sooner or later must figure out that the phrase "less is more" is no paradox. Sizer and his newly established Coalition of Essential Schools are working toward a curriculum which, through attempting less coverage -- by de-emphasizing facts in order to promote thought -- accomplishes much more (**Horace's Compromise**, 109-115).

But we don't need new thinking-skills programs and overpaid consultants to teach us what we already know. English language arts teachers have always been at their best when they move beyond the facts, when they make their students support too-quick responses, when they build on and critically examine a fresh interpretation a student brings up unexpectedly, or when they help a youngster look at an issue from more than one point of view. One hopes that Mr. Cinch (knowledge) has some good moments in his teaching; after all, he may not always teach as he tests. Yet Mrs. Fox (synthesis) and Ms. Edwards (analysis) in all likelihood have more. Their tests are telling signs of instruction in which thinking matters.

So one way teachers can improve the intellectual climate in their classrooms almost overnight is by changing their approaches to measurement. Teacher-made-and-selected tests are frequently flawed, symptomatic of flawed instruction. Higher order thinking is harder to measure. Doing so takes more time, more creativity, and more confidence. And, yes, students frequently don't do as

well analyzing unfamiliar material, such as Ms. Edward's headlines, as they do parroting soon-to-be-forgotten definitions and examples. Thinking hurts. But this is only to be expected; no one does as well on the unfamiliar as he or she might on the commonplace. The obvious point is that lack of success in carrying out analytical tasks provides a powerful case for giving students more opportunity for analysis, not less--even when they fuss about it, as fuss they will.

Thus, as we worry about how external testing programs shape and misshape what we can do in English and language arts classrooms, we mustn't ignore the need to improve our own approaches to evaluation. Moreover, we must be increasingly sensitive to evaluation's effects on what our students learn. When Sergeant Friday, **Dragnet**'s laconic detective still seen on late-night re-runs, asked a witness for "just the facts, ma'am," his television audience knew he could put those facts together into larger patterns and meanings. While many of our students will grow up to be bit players, just witnesses, we must keep Joe Friday in mind. The cybernetic world of the coming century will be led by men and women who learned more than "just the facts." My bet is that Joe Friday knew in his generation and the twenty-first century's leaders will have known in theirs, teachers like Mrs. Fox, Mrs. Barron, and Ms. Edwards. And maybe you.

* * * * * * * * * * *

WORKS CITED

Bloom, Benjamin S. **et al. The Taxonomy of Education Objectives. Handbook 1: Cognitive Domain.** New York: Longmans, Green, 1956.

National Assessment of Educational Progress. **Reading, Thinking and Writing: Results From the 1979-80 National Assessment of Reading and Literature.** Denver: Education Commission of the State, 1981.

Sizer, Theodore. **Horace's Compromise: The Dilemma of the American High School.** Boston: Houghton-Mifflin, 1984.

Large Scale Writing Assessment: An Instructional Message

By
Charles R. Chew

Interest in writing instruction has increased, and we have heard that more has been learned about the writing process—what writers do, what writers should do, and how writers should be taught—in the last ten years than what was learned in the previous hundred years.

We know that students should write and in most cases do not, but we realize that writing is thinking; writing brings the writer into contact with self; writing is a means of learning; writing develops the use of language which in turn assists the student in all subject areas.

With this intensified interest in writing has come an attempt to assess students' writing ability through large scale assessments which focus on direct writing samples. Such assessments are being used locally as well as at the state and national level.

In 1979 when New York State first instituted a writing competency test which required three writing samples from students, it was virtually alone in its attempt to assess students' writing ability through multiple writing samples. Five years have elapsed since the first Regents Competency Test in Writing was given. The program has grown to encompass not only the eleventh grade but the eighth grade and fifth grade as well. Beginning in September of 1985, G.E.D. diploma candidates will be required also to complete satisfactorily a writing sample in order to receive a diploma.

We have heard any number of times that testing can and does drive instruction, and most of us attach a negative response to this idea. If there is a modicum of truth to this, the responsibility placed on the test maker, whether at the local or state level, is an important one. Test makers and those involved in instruction must realize that all tests are "unreal." The test and the test situation can only approximate reality; therefore, direct assessments of writing should, if possible, approximate what we expect when students write.

We are all knowledgeable enough now to recognize that writing is a process—one in which the writer moves through various stages, sometimes in a recursive way. We know that writers draft and redraft and that the concluding stage of the process is often publication or sharing. We know that writers write best about what they know, and such knowledge makes the task of test development difficult. We also know that the process takes time. We seem to have accepted the idea that we gather more information by requiring students to write whole discourses instead of filling in blanks, and have realized that we only gain a limited understanding of the writers' ability through an objective test of discrete skills.

Large Scale Writing Assessment

In New York State, the Writing Test for New York State Elementary Schools administered at grade 5 comes very close to approximating the composing process. Students are required to write two different pieces for the test on two different days. A prewriting section precedes the writing sample and is not evaluated. Students draft a response and then redraft. (Part 2 Writing Task from Spring 1984 follows.)

DIRECTIONS: For Part 2 of the test, you are to write about the topic in the box below.

> This is the beginning of a story.
>
> It was raining when my parents left. I heard thunder and saw lightning. Suddenly the lights in the room began to go off and on. **Then I heard a noise.**
>
> Write about 150 words finishing the story.

NOTES:

1. What noise did you hear? _____
2. What are some things that happened after you heard the noise? _____
3. Think about how you might end the story. _____

In preparing your story, remember to:

1. Write your first draft on scrap paper.
2. Read carefully what you have written.
3. Make any changes that will improve your first draft.
4. Write your final copy of the answer paper given to you by your teacher. Use a pen to write your final copy.
5. Draw a line through any mistake you may make when you are writing your final copy. Make the correction and continue on with your final copy. You do not need to begin a new final copy.
6. Read your final copy before you hand it in to make sure you have not made any copying mistakes.

* * * * * * * * * *

In the tests at the secondary level, the Preliminary Competency Test in Writing at grade 8 requires students to write three pieces, but there is no prewriting connected with the test itself. The same is true for the Regents Competency Test in Writing which is administered in grade 11 and is a requirement for graduation. The Comprehensive Examination in English which is administered usually to average or above average students requires two writing samples. (See chart titled "Types of Writing.")

For grades 5 through 11, the writing tasks are placed within a rhetorical framework which establishes purpose and audience. But in the Comprehensive Examination in English, the student has a choice between a task set within a rhetorical framework and a discrete topic which requires a full rhetorical invention.

Tests, if they are to approximate the reality of what we expect of writers and are to have a positive effect on what happens in the classroom, need to

require many types of writing which writers do. The outline below shows the types of writing assessed in New York State Writing Assessment programs and at what grade level.

Types of Writing

Writing Test for New York State Elementary Schools

Personal Narrative	The writer recounts an experience which she/he had. (My Trip to ____. What My Day Is Like)
Personal Expression	The writer recounts a feeling, an emotion, a reaction. (My Happiest Moment, A Surprise Present)
Description	The writer describes a person, object, or place. (My Favorite Place, The School Cafeteria)
Process Essay	The writer explains how to do something. (How to Take Care of My Pet, How I Learned to Ride a Bike)
Story Starters	The writer completes a story which is started in the writing prompt. (I was alone in the house when suddenly I heard a noise—or saw something)

Preliminary and Regents Competency Test in Writing

Business Letter	The writer at eighth grade writes a letter ordering something. At grade 11, the writer composes a letter of complaint and suggests how to remedy the situation. (A faulty product is under warranty, returned for repair, but repairs are unsatisfactory.)
Report	The writer takes data supplied and prepares a report for another person or the class. (You are preparing a report on Canada. From the notes you have taken, prepare the report for your social studies class.)
Persuasive Discourse	The writer attempts to persuade the reader to take some action by stating the action to be taken and giving reasons why such action should be taken. After the situation has been established, the writer must decide which of two punishments is more appropriate and persuade readers of the school newspaper to agree.

Comprehensive Examination in English

Part II-Essay	The writer, using literature which has been read, responds to a given question which is generic in the sense that a wide variety of literature could be used to respond. (People sometimes act in ways that meet with disapproval from their families, friends, or acquaintances. From the novels, plays, biographies, or autobiographies you have read, choose two in which a character's actions meet with disapproval. For each work, identify the character. Using specific references, describe the actions of the character which are not approved and discuss

fully the reasons for disapproval.) (Give titles and authors.)

Part III-Composition — The writer can choose to do one question from among eight. Two of these are situations which provide a purpose and audience. The other six are discrete topics which require a full rhetorical invention by the writer. (Situation -- Some people show courage without ever having been in battle. You have been assigned to write an article for your school newspaper about an individual you admire because of his or her courage. In an article of about 250-300 words, tell about this person by describing the courageous behavior you admire. Other discrete topics -- Keepsakes, Flea-market Mania, Surprises in Science)

The implications for instruction from such a testing program are positive in many ways. From the material just presented, we realize that students need to write in a variety of forms, for different purposes, and for various audiences. And we should infer that attention must be given to the composing process in all of its stages.

The method of evaluation used in this testing program also speaks to instruction. Students' writing samples are evaluated holistically at grade 5, and modified holistic scoring is used with all the other tests. Because of this rating procedure, a message is delivered to anyone in the state who has any involvement with the testing program. The idea that the whole piece of writing may be worth more than any one single aspect of it is an important message to teachers who have for years spent an inordinate amount of time "red-penciling" errors in students' work.

A more important message is delivered across grade levels, however, if we examine the criteria used to evaluate students' writing in each of the tests discussed. (See chart below.)

Common Elements in Criteria Lists Used for Evaluating Writing Tests *

Elementary	PCT and RCT	English Regents	G.E.D.
Perception of task	Perception of task	Perception of task	To be determined
Plan of organization	Plan of organization	Plan of organization	
Developed with coherence	Developed with coherence and transition	Adequately developed	
Supporting material	Supporting material	Accurate supporting material	
Sense of audience and purpose	Sense of audience and purpose	Sense of purpose implied	
Sentence variety	Sentence structure	Vivid and interesting language (style and diction)	
Mechanics of language	Vocabulary and mechanics	Conventions of standard written English	
	Consistent point of view	Writes in the required form	
	Consistent temporal point of view	Specifies titles, authors and genre	

*Chart prepared by Dorothy B. Hammond, Associate in English Education. Criteria abbreviated for chart use.

BEST COPY AVAILABLE

The fact that seven criteria are virtually the same for grades 5-11 indicates that these elements are seen as essential in a competent piece of writing. The fact that the criterion which focuses on mechanics is not at the top of the list reminds teachers that mechanics, although important, are not the "be all and end all" of written discourse. The fact that the tests are unlimited in time and that length is merely suggested delivers more important messages about the teaching of writing in general.

Other positive aspects of this assessment program include: inservice programs conducted throughout the state to train local educators not only to rate the tests but to develop instructional strategies to meet students' needs in writing, and instructional material developed to supplement the test, and mandated instruction in writing required for those students who fall below the State Reference Point on the writing tests.

Inservice Programs

To meet the needs of educators at the local school level to rate the tests and to devise instructional strategies to meet student needs in writing, the Bureau of English and Reading Education developed a two-year inservice program. The first phase of the program identified fifty key teachers or supervisors which represented geographic areas of the state. These fifty came to Albany and underwent a two and a half day intensive training program. This program focused on rating procedures, questions and answers related to the test, developing a workshop agenda, and actually simulating the role of workshop leader. When these fifty people returned to their local areas, they in turn trained teachers from the local schools who were involved directly with students affected by the writing tests. The success of the program was confirmed by the evaluations done by workshop participants. The sampling by the Bureau of ratings of test papers done locally attested to the reliability of local rating.

The second year of inservice focused on instructional strategies to meet student needs in writing. Basically the same format of training was used except the training sessions at the state and local level focused on analyzing students' written work to ascertain needs, developing instructional approaches to meet these needs, and formulating characteristics of a sound instructional program in writing. This round of inservice also proved successful.

The positive aspects of the program included training over 7,000 local teachers, promulgating information about the writing process, evaluation and instruction, and developing a network of knowledgeable trainers (the original fifty brought to Albany) who could assist us through the state.

Instructional Materials

With all tests, the Bureau develops administrative manuals which serve as invaluable resource materials. These materials include a comprehensive discussion of the writing process, characteristics of a sound writing program, rating procedures, ways to analyze students' writing, and suggested instructional strategies to aid the classroom teacher. The Bureau also published two comprehensive publications, **Helping Student Writers, K-6** and **Helping Student Writers, 7-12**. These publications further explicate the materials described earlier in this paragraph.

Mandated Writing Instruction

Any student who falls below the State Reference Point on the writing tests

must receive additional or remedial instruction. This is not a local option but is required by state law. Parents must be notified not only of the student's grade but must be informed of the remedial program established for the student. These programs must begin no later than one semester after the administration of the test. Students can be removed from remediation if it can be documented that deficiencies have been overcome. At the senior high school level, students must pass the Regents Competency Test in Writing in order to receive a diploma.

However, lest I be accused of being unable to see the negative side of writing assessment, let me suggest some problems, questions, and concerns which need to be addressed by test makers and others interested in improving students' writing ability.

Concerns for Test Makers

1. Samples of students' writing for evaluation and instructional purposes must be obtained throughout the school year **not** only at test time.

2. Research needs to focus on the development of writers over time.

3. Research needs to determine if skills differ appreciably for various types of writing in a test situation.

4. Research needs to ascertain the relationship between writing done during a test and that done by the student at other times.

5. Writing prompts may not tap the experience of the writers.

6. Instruction can be limited to test items. Students may spend an inordinate amount of time writing business letters and structured responses to literature.

7. Evaluators using holistic scoring may not appreciate the fact that more must be done with student papers to plan instruction.

8. Once common elements of competent writing have been identified, instructional strategies need to be developed which will enable teachers to focus on these elements in a total language approach.

9. Inservice connected to a test may be limited when compared to extensive needs of teachers.

Those of us involved in the assessment of writing know how much more we know today than we knew just a few years ago, but there is still much to be learned. Anyone who doubts this needs only to read student test papers for several hours or to examine some of the tests which are being used in the assessment process.

Self-Assessment And The Mastery Of Writing

By
Edgar H. Thompson

In our rush to test students these days, we frequently forget that assessment should provide information for students, as well as for teachers and administrators. According to Lev Vygotsky, only when students are able to detach themselves from the creative task at hand and view their efforts objectively are they in a position to achieve mastery over any cognitive task they are engaged in. This is especially true of writing. After all, when students write something, someone has to make a judgment about its effectiveness. Why not let this "someone" increasingly become the student?

To help my students learn how to assess their own writing and to begin to achieve the kind of mastery suggested by Vygotsky (1962), over the years I have adapted and extended a self-evaluation strategy first developed by Mary Beavens (1977). At the beginning of a course, I give my students a self-evaluation handout listing six items that I want them to address for each major paper they write. (I have taken Items 1 and 6 directly from Beavens' strategy, while I have adapted Items 4 and 5 slightly from what Beavens suggests. Items 2 and 3 are my own additions.) I tell my students to write the self-evaluation as an informal essay or a journal entry.

What follows is the handout that I give to students. Next, I briefly discuss the rationale for each item and then describe how the self-evaluation process works in practice.

Self-Evaluation of Papers

Item 1

How much time did you spend on this paper (including all prewriting)?

Item 2

Describe the process you went through to create this paper by addressing the following questions: Where did you get the idea for the paper? What invention strategy (either your own or one I've given you) did you use to help you explore your subject and get words down on paper? What problems did you have while writing the first draft? What kind of revision strategies did you use to shape and to refine your draft?

Self Assessment for Writing

Item 3

Group Comments: First, list a point your group made about your writing. Then respond to their comments. For example, do you agree or disagree with what they said? What did you do as a result of their feedback? Follow this procedure for at least two or three of the points your group made and arrange your responses in the following manner:

A. Group Comment:
   Your Response:

B. Group Comment:
   Your Response:

Item 4

What are the strengths of your paper? What points still cause you to feel uneasy?

Item 5

What do you want me to look for when I evaluate this paper? What questions do you have for me?

Item 6

What grade would you put on this paper and why?

Rationale for Self-Evaluation Questions

I think it is important that students claim the amount of time they have put into a paper. They need to recognize that good writing takes a lot of effort. Also, the amount of time students say that they have spent on papers gives me a quick reading on how efficient each particular student's writing process is. For example, if some students are spending an inordinate amount of time creating papers and yet the quality of their writing is lacking, then I need to work more intensively with them.

As students respond to Item 2, both the students and I gain insight into how the initial draft evolved and what factors might be causing students to struggle. Frequently, as students address this item, they recognize that had they proceeded in another direction, their writing might have gone more smoothly. I, in turn, learn what kinds of issues I need to be talking about with individual students regarding their writing.

Item 3 reaffirms that students are in control of their own writing. They may receive feedback from several sources, but it is up to students to decide whether or not they will accept or reject advice they are given. Feedback is important, especially feedback from peer groups. To reinforce this idea, when I notice that the comments a group made are similar to the ones I will make in my summary comments to students about their papers, I write a little note on the self-evaluation and say, "Your group was right on target."

Students have an opportunity when dealing with Item 4 to claim for themselves those things that they think they have done well in a paper. Claiming these successes allows students to recognize that they are capable of doing several things well, even though everything in their papers may not be working.

Also, students begin to realize that they are competent to make perceptive judgments about their writing. As they identify problems in their writing and share them with me, I have a chance to address their concerns directly and specifically.

As the course progresses and students learn what factors constitute good writing, Item 5 gives students the chance to both practice and to share what they have learned as their critical judgment sharpens. This item also gives students an opportunity to ask me specific questions which I can answer as I read their self-evaluations. As they ask questions and I answer them, the quality of our communication improves. Students begin to think of me not so much as "the examiner" but as a person who is concerned about helping them improve their writing.

Finally, giving students an opportunity on Item 6 to give themselves a grade forces them to wrestle with the difficulty of summarizing the value of their writing with one letter. Most students find it hard to grade their own work, but they do begin to understand the struggle I face when I have to put a grade on their work. As they justify the grade they have chosen, they learn that there are reasons for assigning a certain grade. I have found, that after the first two or three papers, students are seldom more than one-half letter grade off from the grades I have decided to put on their papers. In fact, after some initial training, I have successfully shared the grading of papers with students, letting their grades and mine count equally.

Teacher Response to Student Self-Evaluations

Reading self-evaluations in addition to papers does add extra work. I find, though, that I can read students' self-evaluations relatively quickly and that I enjoy learning in what light they see their own work. It's more fun than drudgery.

More important, after the first few papers, the effort I expend in evaluation actually lessens. At the beginning of a course, I first read the students' papers and then write my response at the end of their self-evaluations. After I have already written my response, I then read the self-evaluations, making additional comments to help them sharpen their assessment of their own writing. As students become more comfortable and proficient at doing self-evaluations, I read their papers and then go directly to their self-evaluations. I quickly find that the only thing I have to write on their papers is a statement like, "I agree with your assessment," offer praise, and perhaps make a specific suggestion or two. In actuality, students take over the major responsibility of evaluating their own work. I am increasingly able to function as a "sounding board" and a resource person.

Changes in Students' Perceptions of Their Writing

Following is an excerpt from a self-evaluation written by one of my students. This self-evaluation chronicles her reaction to an initial draft of the first paper she wrote in my course about her experiences caring for her horse, working at the stable, and learning some things about responsibility.

Excerpt from First Self-Evaluation

3. Group Feedback
   Group Comment 1: Try to put more description into the main body of the paper, instead of a paragraph at the beginning.

Student Response: I agreed with this before the paper was presented to the group, but I was unsure how to go about it. The group suggested I insert various sentences into the feelings to balance the work.

Group Comment 2: Reword the introduction to include more about the lessons learned at the barn and less description.

Student Response: After rereading the section, I realized the introductory paragraph did just jump into the topic and had a thesis tacked on the end. Previously, the beginning and end was me standing, looking at the structure and recalling past experiences. The paper needed a real conclusion and introduction because this flows better than the technique I had tried.

Group Comment 3: The final, overall concensus about the paper was the need for description throughout and to stress how and why the place was important to me.

Student Response: During revisions, I attempted to follow the advice and create a smoother piece.

4. This paper does contain several strengths and weaknesses. The description of the barn itself was complete enough for the reader to formulate an image, but not overbearing in relation to the rest of the contents. The details and examples on how I learned of maturity and responsibility were also well developed. However, I'm still uneasy about the paper as a whole. My father read it and said it seemed all right, but it just didn't convey the things I intended. Usually, I don't have problems transferring ideas to paper, but writing this paper was almost a chore.

Though this student made an honest attempt to evaluate her writing, two things were working against her at this early point in the course. First, her group's comments were accurate, but they were vague. Her peers were not yet skilled at making specific comments. Second, she recognized that the paper was not communicating what she wanted, but she missed the real reason why. She thought her examples and details were sufficient. In fact, they were not. Still, her comments on this early self-evaluation allowed me to be pointed in my summary comment to her. Following is what I wrote:

You've got an excellent subject for a paper, but you need to do some revision.

1. Your introduction doesn't add much to the paper as it is. Why not start with the second paragraph?

2. Describe in detail some of the experiences that happened to you there. For example,
   a. the shock of having to take responsibility,
   b. dealing with the crude ways and habits of adults, and
   c. gaining the respect of those older than you.

   Put us, with specific details, into each of these experiences.

3. In an analytical summary (two paragraphs) tie together how each of these experiences helped you mature.

With revision, excellent potential.

She gave her paper a B-/C+ and I gave it a C-. Notice that I had to carefully analyze the problems in the paper and make very specific suggestions, which took some effort on my part.

After further feedback from her peers and two conferences with me, she revised this paper three months after she had submitted the first draft. Following is an excerpt from the self-evaluation of her revision:

Excerpt from Second Self-Evaluation

3. Group Feedback
   Comment 1: Make a clearer definition of the actual timeframe.

   Student Response: Okay, I'm not sure why that is so important, but I will.

   Comment 2: You need to expand and personalize the section with the vet.

   Student Response: I agree that it should be more than a sentence description. I'll try.

   Comment 3: Continue the incident with the drunk horse owner and tell us what happened.

   Student Response: Yes, I did leave that hanging and unrelated to the main idea.

4. After rereading the very first version of this paper, I can definitely see vast improvements. The paper is more explanatory and personal now. By using specific incidents and first person, I tried to put the reader into the situation. I know that there are problem areas even still, but I'm not sure what they are—maybe just polishing.

My response to her revision and self-evaluation was as follows:

> I think you have made very solid improvements in this paper. Generally, you comment on them in your self-evaluation, so there is no need for me to be repetitious. You might work on sharpening the detail in the first paragraph, but I think that's minor at this point. You've made great strides in your maturity as a writer in this paper.

My grade for this paper was the same as hers, B+. Also, notice the differences in this self-evaluation from the first one. First, comments from her peers were specific enough so that she could act on their suggestions. Second, she was confident that she knew what was and was not working in this paper. As a result of these changes in her self-evaluation, I was able to make my summary comment more general. Essentially, all I had to do was to assure her that her own assessment was accurate and to praise her efforts. My comment on this second paper took much less effort than did my comment on her first paper. The energy

involved in evaluating the paper shifted from me to the student, where it should be.

At the beginning of the course, this student had a hard time accepting criticism about her work. As the course progressed, however, she began to listen to feedback that she was being given and gained confidence that she could improve her writing by stepping back and evaluating it objectively. Though not the only factor, I believe that the self-evaluation process stimulated and reinforced this change.

When we assess our students' work, our major goal is usually to gather information about their performance to help us help them achieve mastery over the tasks they are asked to do. However, it seems to me that it is also in the best interest of our students to teach them how to effectively evaluate their own work. (Our job is made easier.) Through objective self-evaluation and the interactions that take place between students and teachers during this process, students begin to approach the level of mastery we want them to achieve.

* * * * * * * * * *

Works Cited

Beaven, Mary H. "Individualized Goal Setting, Self-Evaluation, and Peer Evaluation." In **Evaluating Writing.** Ed. Charles R. Cooper and Lee Odell. Urbana, Illinois: NCTE, 1977: 135-156.

Vygotsky, Lev. **Thought and Language.** Trans. and Ed. Eugenia Hanfmann and Gertrude Vakar. Cambridge, Massachusetts: The M.I.T. Press, 1962.

Trends And Issues In The Use And Misuse Of Oral Language Proficiency Testing For Bilingual Program Planning

By
Josefina V. Tinajero

The 1974 Lau vs Nichols decision of the United States Supreme Court, as well as recent federal statutes and state legislation, require that school districts develop and implement appropriate programs designed to provide a linguistically comprehensible education for the more than 3.6 million limited English proficient (LEP) students in our schools today. The Lau vs Nichols decision states in part that

> Where inability to speak and understand the English language excludes national origin-minority group children from effective participation in the educational program offered by a school district, the district must take affirmative steps to rectify the language deficiency in order to open its instructional program to these children. (414 U.S. 563.1974)

This Supreme Court ruling, together with Title VII of the Elementary and Secondary Education Act and legislation from twelve states which currently mandate bilingual education, have given rise to the development and use of language proficiency tests. A number of these tests have recently been developed and are currently used to identify students in need of special linguistic and instructional services. Furthermore, decisions concerning appropriate services for these children, particularly as they relate to the language arts program, have been based primarily on the results obtained from these tests.

This paper begins by providing a brief background of the procedures and instruments used for the identification and assessment of LEP students. Next, it discusses how the results obtained from oral language proficiency tests are currently used for planning and designing the language arts program for LEP students. And finally, it discusses some of the problems that exist as a result of misusing these tests.

Student Identification and Placement—A Three Step Process

Reyes (1980), from California's State Department of Education, maintains that proper identification of LEP students is essential to the development of appropriate treatments for the acquisition of standard English by those students. Furthermore, Cummins (1981) maintains that proper identification and placement is equally essential to the development of appropriate treatments

which provide for the conceptual development and academic achievement of students, particularly as those treatments relate to the use of the students' home language.

Because proper identification and placement of students plays a vital role in providing appropriate instructional treatments for them, most states have prescribed a three-step process to assist school districts in identifying students and in designing viable instructional programs for them. First, a survey of the home language is conducted to identify students who are potentially limited-English or non-English speaking. Next, an oral language proficiency test is administered to the identified students in order to determine the strengths and weaknesses of their English language skills. Proficient English speakers are then eliminated from further testing while students identified as LEP are tested further in their home language to determine their proficiency in that language. Figure 1 summarizes this identification and placement process.

Language Assessment Tests

A number of assessment instruments are currently used by school districts throughout the United States to assess students' language proficiency and thereby to identify and place them in appropriate educational programs. The focus of these instructional programs is determined primarily from the results obtained from these tests. A survey of several states including New York, New Mexico, California, and Texas indicated that four assessment tests are commonly used by school districts. These tests are: (1) **Oral Language Dominance Measure - OLDM**, also known as the **PAL**; (2) **Basic Inventory of Natural Language - BINL**; (3) Bilingual **Syntax Measure - BSM**; and, (4) **Language Assessment Scales - LAS.**

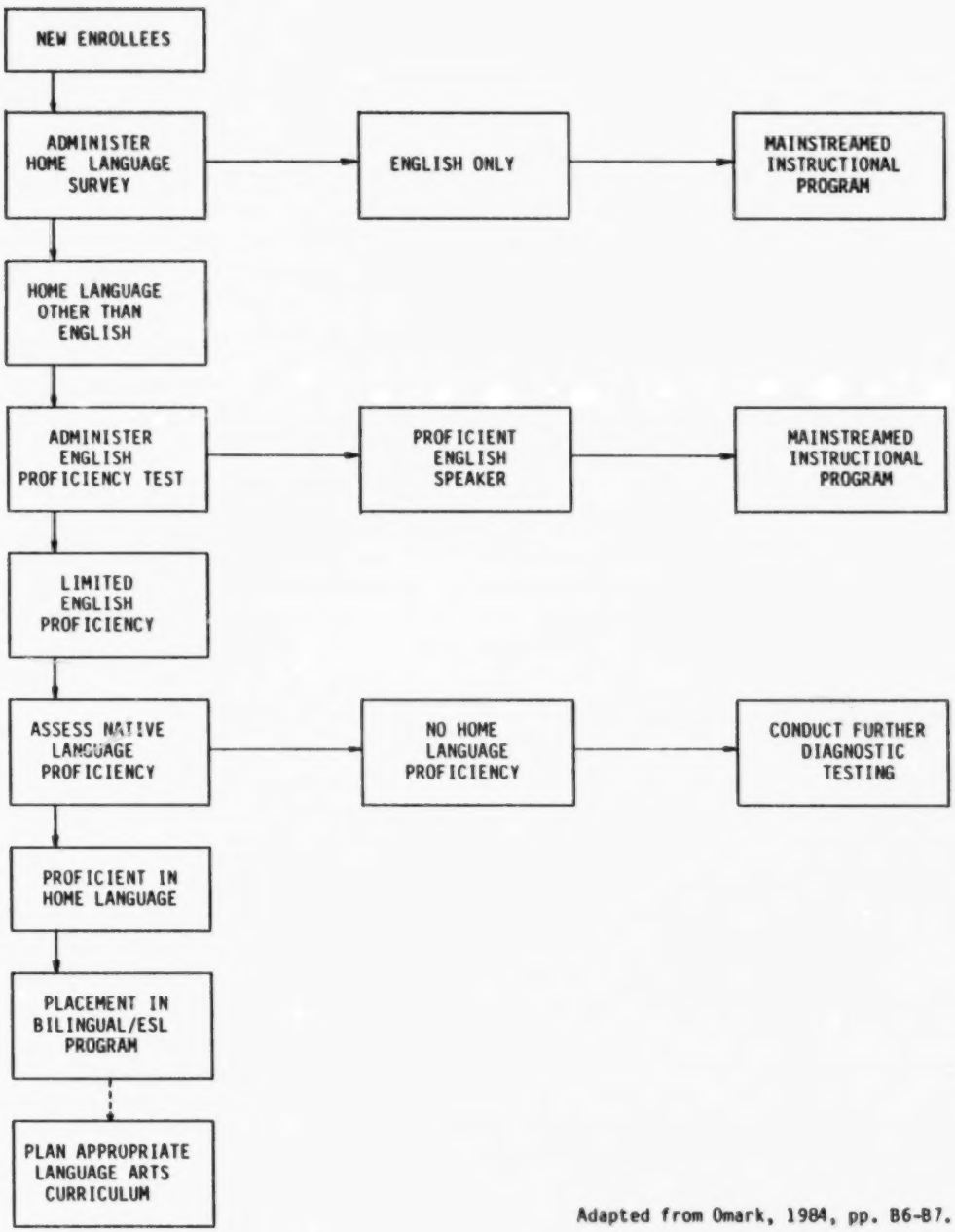
Major Linguistic Structures Measured

Each of these language assessment tests has been designed to measure one or more of the linguistic structures of oral language: phonology, lexicon, semantics, and syntax. Fluency is also considered an important measure of oral language proficiency by some of these tests. Furthermore, each test is designed to measure students' use of these linguistic structures by measuring their communication skills in oral production or aural perception or both. Each of the areas assessed by these tests helps to determine "if a student has sufficient receptive and productive command over spoken English and its uses to be able to function in an all-English medium classroom" (Rosenbaum, 1980). Table 1 summarizes the linguistic structures evaluated and the communication skill or skills used as a basis for evaluation by each of the four language assessment tests (Tinajero, 1980).

Determining Language Dominance and Proficiency

Each of the four tests listed in Table 1 has a Spanish and an English version and measures various aspects of language proficiency in each of these languages. Each test also yields an independent measure of proficiency ranging from little or no proficiency in the language (level 1) to native-like fluency (levels 4, 5, or 6 in most tests). These independent language proficiency scores are then used to design appropriate instruction in each component of the language arts (listening, speaking, reading, writing).

FIGURE 1
STUDENT IDENTIFICATION AND PLACEMENT PROCESS



Adapted from Omark, 1984, pp. B6-B7.

TABLE 1
LINGUISTIC STRUCTURES AND COMMUNICATION SKILLS
ASSESSED BY SIX LANGUAGE ASSESSMENT TESTS

Oral Language Assessment Tests	Linguistic Structures								
	Phonology		Lexicon		Semantics		Syntax		Fluency
	Communication Skills								
	Pro-duction	Per-ception	Pro-duction	Per-ception	Pro-duction	Per-ception	Pro-duction	Per-ception	Pro-duction
PAL (OLDM)							X		X
BINL							X		X
LAS	X	X	X			X	X		X
BSM							X		

The results obtained from these tests also yield important information about students' strengths and weaknesses in each of the linguistic structures and communication skills measured. The areas of deficiency that are found also become important pieces of information for planning the language arts curriculum content.

Furthermore, a comparison of students' proficiency scores in Spanish or in any other home language, and in English yields their stronger language and their relative proficiency classification (Dulay and Burt, 1980). This information allows school districts to classify students into one of the following categories and serves as a guide for planning appropriate corresponding instructional programs.

1. Monolingual speaker of the language other than English
2. Predominantly speaker of the language other than English
3. Bilingual
4. Predominantly speaker of English
5. Monolingual speaker of English.

Table 2, from the **Oral Language Dominance Measure Examiner's Manual** (1978), is an example of how the independent proficiency scores for each of the two languages are used to determine the relative proficiency classification.

BEST COPY AVAILABLE

TABLE 2
DETERMINING LANGUAGE DOMINANCE

Language Levels			Language Levels			Language Levels			Language Levels			Language Levels		
Eng.	Span.	Dom. Lang.	Eng.	Span.	Dom. Lang.	Eng.	Span.	Dom. Lang.	Eng.	Span.	Dom. Lang.	Eng.	Span.	Dom. Lang.
5	1	Eng.	4	1	Eng.	3	1	Eng.	2	1	**	1	1	**
5	2	Eng.	4	2	Eng.	3	2	*	2	2	**	1	2	**
5	3	Eng.	4	3	*	3	3	*	2	3	Span.	1	3	Span.
5	4	Eng.	4	4	*	3	4	Span.	2	4	Span.	1	4	Span.
5	5	Bil.	4	5	Span.	3	5	Span.	2	5	Span.	1	5	Span.

*Begin reading instruction in the child's mother tongue.
**Scores extremely low for both languages. Please discuss case with your principal and consultant.

Source: **OLOM Examiner's Manual**, 1974

Once their relative proficiency classification has been determined, students are placed in appropriate educational programs. Generally, students classified as 1's (monolingual speakers of the language other than English) and 2's (predominantly speakers of the language other than English) are classified as LEP and are therefore assigned to bilingual or ESL classes. It is important to note, however, that although all these students share some characteristics in their use of the English language, they also show a variety of levels of proficiency in their acquisition of English. Thus, although they all have the label LEP, "we are realistically dealing with children who represent a wide latitude of linguistic behaviors instead of a single group alike in their non-use of English" (Massachusetts State Department of Education, 1978). These linguistic behaviors and their corresponding program designs are discussed next.

Application of Language Proficiency Test Results in Planning the Language Arts Program

A sound bilingual education program for LEP students generally consists of the following six components:

1. Basic concepts are taught in students' dominant language
2. Language development is provided in students' dominant language
3. Language development is provided in the English language
4. Subject matter and concepts are taught in students' dominant language
5. Subject matter and concepts are taught in English
6. Specific attention is given to instilling in students a positive identification with their cultural heritage.

BEST COPY AVAILABLE

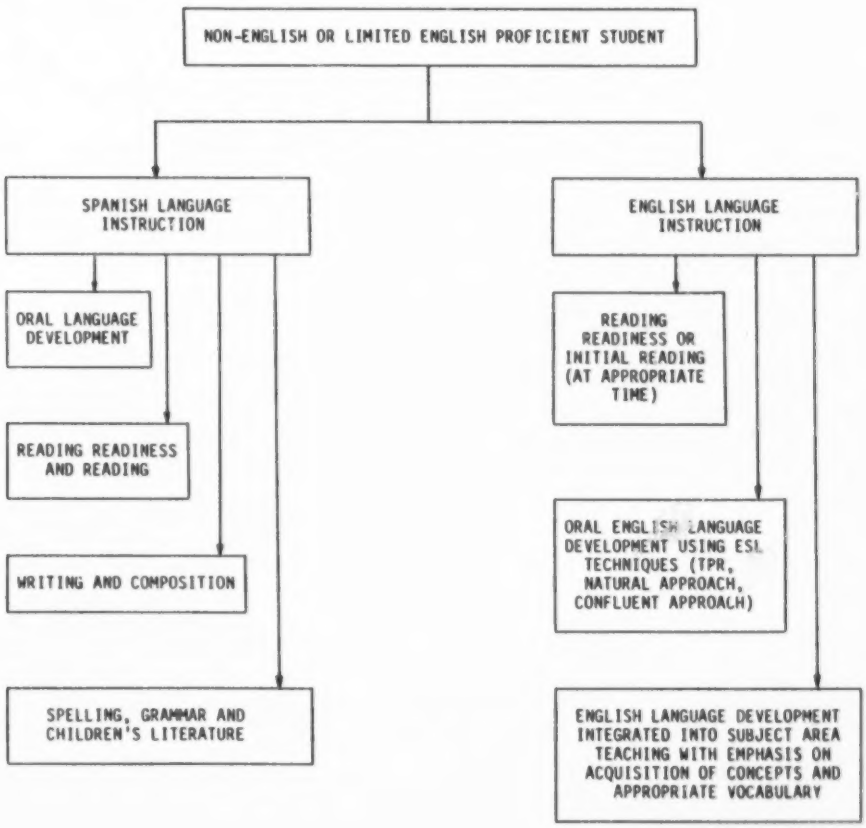
The first three of these six components function as a framework for developing the language arts program particularly as they relate to language use within each component.

Conceptual Framework for Curriculum Content and Language Use

The home language and English are both used to the extent necessary to achieve two educational objectives for bilingual students: (1) to improve their English language skills and (2) to enhance their academic achievement. Figure 2 shows the features of the language arts program which have generally been implemented by school districts in an effort to accomplish these two objectives.

As shown in Figure 2, the home language is used initially for instruction in all areas of the language arts except for English oral language development. As the students acquire higher levels of English language proficiency, reading readiness and formal reading instruction are also conducted in English. Complete transfer to English instruction within all components of the language arts is generally accomplished by most school districts within three years. Again, language assessment plays a central role in determining students' readiness to begin instruction in English. Language assessment also plays a central role in determining when students have acquired sufficient English language skills to function exclusively in English.

FIGURE 2
CONCEPTUAL FRAMEWORK FOR LANGUAGE ARTS CURRICULUM CONTENT


NON-ENGLISH OR LIMITED ENGLISH PROFICIENT STUDENT
SPANISH LANGUAGE INSTRUCTION
ENGLISH LANGUAGE INSTRUCTION
ORAL LANGUAGE DEVELOPMENT
READING READINESS AND READING
WRITING AND COMPOSITION
SPELLING, GRAMMAR AND CHILDREN'S LITERATURE
READING READINESS OR INITIAL READING (AT APPROPRIATE TIME)
ORAL ENGLISH LANGUAGE DEVELOPMENT USING ESL TECHNIQUES (TPR, NATURAL APPROACH, CONFLUENT APPROACH)
ENGLISH LANGUAGE DEVELOPMENT INTEGRATED INTO SUBJECT AREA TEACHING WITH EMPHASIS ON ACQUISITION OF CONCEPTS AND APPROPRIATE VOCABULARY

Level of Proficiency and English Language Development

The ESL (English as a Second Language) component of the language arts program is designed to provide students with special English language instruction. Students are assigned to appropriate levels of instruction depending on their English language proficiency as determined by an assessment test. For example, the Texas Education Agency has identified three categories of English language proficiency and has prescribed an appropriate curriculum design for each category. (See Table 3.) School districts use this framework to develop ESL activities for LEP students.

TABLE 3
ESL IN THE STATE PLAN

Categories of English Language Development	Program/Curriculum Design
Level I - Speaks and/or understands little or no English.	Much experience in listening and speaking with special attention to pronunciation, vocabulary, and basic syntactic structures in English.
Level II - Speaks and/or understands some English.	Continue development of listening and speaking skills along with increasing attention to the skills of reading and writing. Emphasis is on reading comprehension and on writing English sentences.
Level III - Speaks and/or understands English but has some difficulty performing ordinary classwork in English.	Equal emphasis on all language skills. Activities focus on extensive practice in comprehending and producing a variety of oral and written forms in English.

Source: Texas State Plan for Bilingual Education, 1978, p. 11.

In California, a three-stage language proficiency classification has also been recommended as a basis for designing the ESL program. These three stages are: (1) the pre-production stage; (2) the early production stage; and (3) the speech emergence stage. These stages are then used to group children for appropriate instruction and to select appropriate methodologies. Table 4 summarizes the methodologies recommended by Calderon (1985) for each stage of language proficiency. A brief description of each stage is also included.

COPY AVAILABLE

TABLE 4
THREE STAGES OF LANGUAGE PROFICIENCY AND RECOMMENDED CURRICULUM DESIGN

Stage	Description	Approach
Pre Production	The silent period—a time for receptive language acquisition lasting an average of ten hours to three months of instruction. Students develop literary skills.	Total Physical Response (TPR) Approach Natural Approach
Early Production	Period of limited production—progress will vary lasting an average of six months to two years of instruction.	Total Physical Response (TPR) Approach Natural Approach (oral language development integrated with the teaching of subject matter content)
Speech Emergence	Period of expanded production—progress will vary. Error correction is limited to modeling and expansion of student's language. This period may last from two to four years.	Total Physical Response (TPR) Approach Natural Approach Confluent Approach (oral language development integrated with the teaching of subject matter content)

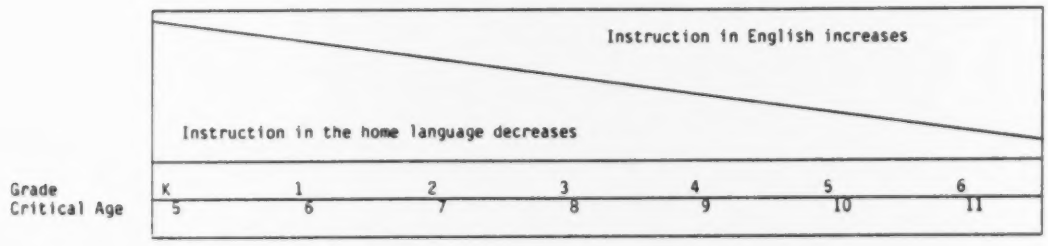
Source: Calderon, 1985

Curriculum Design for Transfer and Test Results

Federally funded and state bilingual programs are transitional in nature. This means that students' participation in these programs is generally limited to three years or "until such time as the student achieves a level of English language skill which will enable him [or her] to perform successfully in classes in which instruction is given only in English, whichever shall first occur" (Medina-Seidner, 1980). This also means that bilingual programs must be designed to gradually transfer students from extensive use of the home language to exclusive use of English. Again, language assessment plays a central role in determining which students are ready to advance to higher levels of instruction in English and which ones are ready to transfer out of the special program into regular English-medium classes.

Figure 3 serves as a framework for planning and designing a bilingual transfer program.

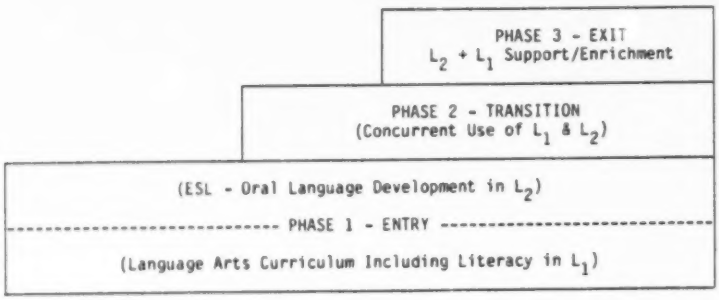
FIGURE 3
FRAMEWORK FOR EFFECTIVE TRANSFER OF
LANGUAGE ARTS CONTENT FROM L1 TO L2



BEST COPY AVAILABLE

This framework has generally been used by school districts for planning their language arts transfer programs. The design of this framework has been guided by three major objectives: (1) to develop high levels of communication skills in both the home language and English; (2) to facilitate and support the gradual transfer of literacy skills and concepts from the home language to English; and (3) to mainstream students into the regular all-English curriculum. To accomplish these objectives, Gonzales-Baker and Tinajero (1984) suggested that the language arts program be divided into three distinct, overlapping phases of bilingual instruction and that students progress through each phase as they acquire higher levels of English proficiency. (See Figure 4.)

FIGURE 4
PHASES OF BILINGUAL INSTRUCTION IN THE LANGUAGE ARTS


PHASE 3 - EXIT
L_2 + L_1 Support/Enrichment
PHASE 2 - TRANSITION
(Concurrent Use of L_1 & L_2)
(ESL - Oral Language Development in L_2)
PHASE 1 - ENTRY
(Language Arts Curriculum Including Literacy in L_1)

Source: Gonzalez-Baker and Tinajero, 1984.

Phase 1. This phase is characterized by extensive use of the home language to teach all basic concepts and to provide instruction in oral language development, reading readiness, reading, spelling, grammar, writing, and children's literature. In addition, students are provided with extensive English oral language development in which a variety of ESL methodologies are used. When students achieve solid academic success in Spanish reading and other language arts areas and develop sufficient English language skills, progress into the Transitional Phase is made. The results obtained from the English version of a proficiency test are used to determine readiness for transfer, particularly as they refer to readiness for beginning formal instruction in English reading. Most school districts require that students score a 4 or 5 on the English proficiency test before beginning English reading (Gonzalez-Baker and Tinajero, 1984).

Phase 2. The Transitional Phase involves the use of both the home language and English to the extent necessary to allow students to participate in academic tasks in all areas of the language arts. Switching between the home language and English is appropriate if both languages are needed to facilitate learning. Regardless of which language is used, instruction is aimed at the development of cognitive, thinking, and problem solving skills. Progress into the next phase is generally determined by a score of 5 on the English proficiency test plus an adequate score (normally at the fortieth percentile or higher) on the language arts portion of an achievement test recommended by the state.

Phase 3. Once a determination has been made that a student can undertake academic tasks in English, the Exit Phase follows. Normally, when students have reached this level of English language proficiency, all instruction in the home language is terminated or maintained for enrichment purposes.

Issues in Language Testing

Issues in the Misuse (And Abuse) of Oral Language Proficiency Tests

Duron (1983) maintains that tests can be used constructively if they are considered aids or tools to be interpreted carefully. As with any other type of tests, the results obtained from oral language proficiency tests have not always been used constructively. Cummins (1979, 1980a, 1980b) has conducted research in language assessment within a bilingual context and maintains that "failure to take account of those limitations of oral language proficiency tests can result in incorrect diagnosis of students' language difficulties and in inappropriate placement" (Cummins, 1980a, p. 79).

Four issues that relate to the misuse of oral language proficiency tests are discussed here: (1) decisions which concern all areas of the language arts program but which are based on measures that only assess oral proficiency; (2) decisions which are based on levels of English proficiency and ignore proficiency in the home language; (3) decisions which are made on tests which only measure surface proficiency in the English language; and (4) the labeling of students as "English proficient" by measures which only assess their surface proficiency of the language; this often results in the placement of a large number of LEP students in special education classes.

Language Tests -- Primarily Measures of Oral Proficiency

Duron (1983) maintains that "a problem exists in that most language proficiency tests deal with oral language only" (p. 69). However, these tests are almost the only ones used for bilingual program design and planning. Decisions concerning placement, transfer, and exit from bilingual programs are all based on measures of oral proficiency only. Linguists generally agree that the tests used for these purposes should be more global in scope and that they should measure reading and writing skills as well as verbal receptive and expressive skills (Duron, p. 67).

English vs Native Language Proficiency

Cummins (1979, 1981) and Skutnabb-Kangas and Toukomaa (1976) have empirically supported the view that there is a strong relationship between the development of the home language and English. Their data suggest that a certain level of development in the home language is necessary for learning English and for achievement in school. Cummins (1979) maintains that "the development of competence in a second language is partially a function of the type of competence already developed in the home language at the time when extensive exposure to the second language begins" (p. 222). He also maintains that measures of students' literacy levels in the native language are also critical factors to be considered when assessing language proficiency and when making decisions concerning transfer and exit. In practice, these literacy levels are rarely considered when determining students' readiness for engaging in academic work in English.

The BICS vs CALP Distinction

Cummins (1979, 1980a, 1980b, 1981) and Skutnabb-Kangas and Toukomaa (1976) have distinguished between two levels of language proficiency—BICS (Basic Interpersonal Communication Skills) and CALP (Cognitive Academic Language Proficiency). BICS refers to the level of language proficiency needed to function in every-day situations. CALP refers to the level of language proficiency

needed to function in academic tasks. Both Cummins and Skutnabb-Kangas and Toukomaa maintain that it takes considerably longer to develop CALP than it takes to develop BICS. Cummins (1979, 1981) has argued that the proficiency tests which are currently used throughout the United States for bilingual program planning primarily assess BICS—face-to-face communication skills. Therefore, students exited on the basis of these tests are likely to experience considerable difficulty in English-only programs which require CALP to complete academic type tasks.

Placement of LEP Students in Special Education Classes

Cummins (1980a, 1982) maintains that more serious consequences exist as a result of school districts not being aware of the distinction between BICS and CALP. He maintains that, if tests measure face-to-face surface proficiency and thereby label children as proficient English speakers, then "presumably with sufficient well developed English skills there is danger that the following line of reasoning is involved when results from psychological assessment tests are analyzed:

> Because language minority students are fluent in English, their poor academic performance and/or test scores cannot be attributed to lack of proficiency in English. Therefore, the student must either have deficient cognitive abilities or be poorly motivated (Cummins, 1983b, p. 3).

Cummins (1983b) concludes by saying that "these misconceptions reflect a failure to recognize the crucial differences between the language proficiency involved in face-to-face communication and that involved in academic tasks" (p. 3).

Summary

The Lau vs Nichols decision of the Supreme Court and federal and state legislation from twelve states that currently mandate bilingual education have given rise to the need for language proficiency assessment. Schools are required to identify and place limited English proficient students in appropriate instructional programs to develop their English language skills and to support their cognitive development. This paper has provided a brief overview of the procedures and instruments generally used by school districts in the United States for identification and placement of these students. It has also discussed how the results obtained from these tests have generally been used to plan and design the language arts program within a bilingual education context. Moreover, four major problems resulting from the misuse of these tests have been discussed.

* * * * * * * * * * *

Works Cited

Calderon, Margarita. Unpublished duplicated handout for workshop participants. San Francisco Unified School District. Multidistrict Teacher Training Institute (MTTI). San Francisco, California, 1985.

Cummins, Jim. "Linguistic Interdependence and the Educational Development of Bilingual Children." **Review of Educational Research** 49 (1979): 222-251.

_____. "Psychological Assessment of Immigrant Children: Logic of Intuition." **Journal of Multilingual and Multicultural Development** 1 (1980a): 97-111.

_____. "The Entry and Exit Fallacy in Bilingual Education." **NABE Journal** 4 (1980b): 25-59.

_____. "The Role of Language Development in Promoting Educational Success of Language Minority Students." **Schooling and Language Minority Students: A Theoretical Framework.** California State Department of Education. Los Angeles: Evaluation Dissemination and Assessment Center, 1981.

_____. "Tests, Achievement, and Bilingual Students." **Focus.** Rosslyn, Virginia: National Clearinghouse for Bilingual Education 1 (1982): 1-7.

_____. "Conceptual and Linguistic Foundations of Language Assessment." **Issues of Language Assessment Volume II: Language Assessment and Curriculum Planning.** Ed. S.S. Seidner. Chicago: Illinois State Board of Education, 1983a. 7-16.

_____. "Wanted: A Theoretical Framework for Relating Language Proficiency to Academic Achievement Among Bilingual Students." **Language Proficiency and Academic Achievement.** Ed. Charlene Rivera. Clevedon, England: Multilingual Matters 1983b. 2-19.

Dulay, Heidi and Burt Marina. "The Relative Proficiency of Limited English Proficient Students." **NABE Journal** 4 (1980): 1-23.

Duron, Susan. "The Use and Misuse of Instruments." **Issues of Language Assessment Volume II: Language Assessment and Curriculum Planning.** Ed. S.S. Seidner. Chicago: Illinois State Board of Education, 1983: 67-72.

Fillmore, Lily W. "Language Minority Students and School Participation: What Kind of English is Needed?" **Journal of Education** 64 (1982) 143-156.

Gonzales-Baker, Maria and Josefina V. Tinajero. "Juanito's Reading Problems Revisited: Setting Criteria for Transfer of Literacy Skills and Academic Content from L_1 to L_2." 13th Annual International Bilingual Bicultural Education Conference of the National Association of Bilingual Education. San Antonio, Texas, April 1984.

Lau vs Nichols, 414 U.S. 563, 566. **The U.S. Law Week Supreme Court Opinions.** Washington, D.C.: The Bureau of National Affairs, 42 (1974): 4165-4168.

Massachusetts Department of Education. **Proceedings: Diagnosis and Intervention in Bilingual Special Education: Searching for New Alternatives.** Massachusetts: Division of Special Education, Bureau of Program Development and Evaluation, 1978.

Medina-Seidner, Maria. "The Assessment of Language Minority Students: Current Trends in the State of Illinois." **Current Issues in Bilingual Education.** Ed. J.E. Alatis. Washington, D.C.: Georgetown University Press, 1980: 253-258.

Omark, Donald R. and Daniel L. Watson. **Assessing Bilingual Exceptional Children: In-Service Manual.** San Diego, California: Los Amigos Research Associates, 1984.

Reyes, Ramiro. "Identification and Assessment of LES/NES Students in California." **Current Issues in Bilingual Education.** Ed. J.E. Alatis. Washington, D.C.: Georgetown University Press, 1980: 247-252.

Rivera, Charlene and Carmen Simich. "Issues in the Assessment of Language Proficiency of Language Minority Students." **NABE Journal** 6 (1981): 19-39.

Rodriquez-Brown, Flora V. "Use of Native Language Tests for Program Planning Purposes." **Issues of Language Assessment Volume II: Language Assessment and Curriculum Planning.** Ed. S.S. Seidner. Chicago: Illinois State Board of Education, 1983: 121-125.

Rosenbaum, Harvey. "The Development and Structure of the Language Skills Framework of the Student Placement System for Bilingual Programs." **Current Issues in Bilingual Education.** Ed. J.E. Alatis. Washington, D.C.: Georgetown University Press, 1980: 31-63.

Shutnabb-Kangas, Tove and Pertti Toukomaa. **Teaching Migrant Children's Mother Tongue and Learning the Language of the Host Country in the Context of the Socio-Cultural Situation of the Migrant Family.** Helsinki: The Finnish National Commission on UNESCO, 1976.

Texas Education Agency. **Texas State Plan for Bilingual Education: A System for Meeting the Needs of Limited English-Speaking Students in Texas Public Schools, Grades K Through 12.** Austin, Texas, 1978.

Tinajero, Josefina V. **Oral Language Competencies Found in Language Assessment or Achievement Tests as Related to Teachers' Perceptions of Their Importance, Time-Treatment, and Received Training.** Diss. Ann Arbor, Michigan: University Microfilms International, 1980.

Tests Cited

Basic Inventory of Natural Language. San Bernardino, California: Checkpoint Systems, 1978.

De Avila, E.A. and Sharon E. Duncan. **Language Assessment Scales.** Corte Madera, California: Linguametrics Group, Inc., 1978.

Burt, Marina K., Heidi C. Dulay, and Eduardo Hernandez. **Bilingual Syntax Measure.** New York: Harcourt, Brace, Jovanovich, Inc., 1975.

El Paso Public Schools. **Oral Language Dominance Measure.** El Paso, Texas: Department of Curriculum and Instruction, 1978.

Appendix

In the fall of 1983, representatives from the Michigan Association of Speech Communication (MASC), the Michigan Council of Teachers of English (MCTE), and the Michigan Reading Association (MRA) joined together to form the Literacy Consortium. The consortium began to address current issues of mutual concern in advancing literacy in Michigan. The literacy statement on testing and assessment is a result of deliberations started at the Gull Lake Conference in 1984, a two day meeting supported by a grant from the Michigan Council for the Humanities.

Literacy Statement On Testing And Assessment

The measurement of learning is extremely complex. Because perceptions of this complexity vary among and within lay and professional groups, testing as a measurement of learning has definite limitations. Because of these limitations, abuses of testing have occurred:

* Decision-making about students and school programs relies too heavily on formal test results.

* Too much instructional time is used for test-taking and for the preparation, administration, scoring and interpretation of tests.

* Goals for instruction are limited because of the narrow range of language competencies measured by the tests.

* Simple testing formats are used for measuring complex language competencies. These tests usually measure what students will or can do in a testing situation, not what they will or can do in real-life situations.

* A large portion of limited funds for instruction is used for tests.

Therefore, to promote literacy, we recommend three testing and assessment proposals.

A. School district policies on testing should be locally developed.

B. Both teacher-developed measures and standardized assessment devices should be used to make curriculum decisions to assess student progress, and to assign students to particular classes.

C. The use of test results should not be the sole criterion for program funding.

Contributors

James M. Brewbaker *is Professor of English Education and Chair, Department of Curriculum and Instruction at Columbus College. Twice President of the Georgia Council of Teachers of English, he has published in* ***English Journal****, the* ***ALAN Review****, and* ***The Journal of Teaching Writing****.*

Charles R. Chew *is Chief of the Bureau of English and Reading Education for New York State and develops with his staff examinations pertaining to English and Reading, English/Language Arts curriculum, and monitors Chapter I programs in reading and writing.*

Mary Jane Curry *teaches adolescents with learning disabilities. She has also taught at Stillman College, Tuscaloosa, where she helped set up its Writing Lab and designed a remedial composition curriculum. She holds a master's degree in English (University of Alabama, 1975) and one in learning disabilities (U.A.B., 1982).*

Sheila Fitzgerald *is a Professor of the Department of Teacher Education at Michigan State University with responsibility for instruction in methods of teaching Language Arts. She is the current President of the National Council of Teachers of English and a Past President of the Michigan Council of Teachers of English.*

George F. Madaus *is Director of the Center for the Study of Testing, Evaluation, and Educational Policy at Boston College, Massachusetts, and a former President of the National Council on Measurement and Evaluation. A different version of his article recently appeared in* ***Phi Delta Kappan*** *magazine.*

Edgar H. Thompson *is an Assistant Professor of Education and Director of Elementary Education at Emory and Henry College, Emory, VA. He has several years of high school and college English teaching experience.*

Josefina Villamil Tinajero *is an Assistant Professor at the University of Texas at El Paso. Recent publications include co-author of a 1-5 Spanish supplementary reading series and a 1-5 basal reading series both being published by MacMillan Publishing Company. She was recently named an outstanding Hispanic leader by the Mexican American National Association.*

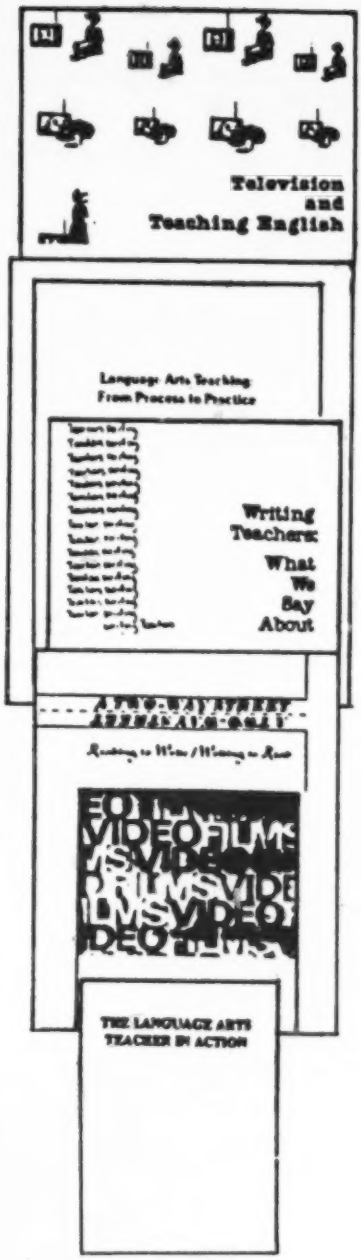
Marilyn Wilson *is an Assistant Professor in the Department of English at Michigan State University, where she teaches courses in reading, applied linguistics, and English education. She has made presentations at MCTE, NCTE, IRA, and TESOL, and has published in* ***The Reading Teacher*** *and in the* ***NCTE International Assembly Newsletter.***

Petey Young *is an Associate Professor in the Department of Education, Southern Oregon State College, where she teaches undergraduate and graduate courses in reading. Previously, she specialized in test construction in the English Department of the University of Ethiopia. Last summer she conducted a workshop for teachers in Gansu Province, People's Republic of China.*

* * * * * * * * * * *

Cover Design by: ***David Barlow****, Kendall School of Design, Grand Rapids, Michigan*

Editorial Assistance from: ***Kathy Drzick****, Western Michigan University*
Jill Rae Van Antwerp*, Lowell High School*


Television and Teaching English
Language Arts Teaching From Process to Practice
Writing Teachers: What We Say About
A Two-Way Street
Reading to Write / Writing to Read
VIDEO FILMS
THE LANGUAGE ARTS TEACHER IN ACTION

☆Television and Teaching English

Nine articles of practical ideas about how television can be used to enrich the teaching of English: teaching critical thinking skills through television; using television as a motivator in teaching; and teaching literature through the use of television are presented with suggestions for the elementary, middle school, and secondary teacher. 40 pages. Edited by Rhoda Maxwell.

☆Language Arts Teaching: From Process to Practice

Nine articles emphasizing the language arts teacher as participator: writing with students, reading aloud to students, and exploring new literature for students. Articles by John Woods, Herbert Scott, and Stephen Dunning on teaching poetry to, and writing poetry with, your students. Using historical fiction (with annotated bibliography). And helping young children develop their visual vocabulary: an analysis of a quality children's book. 79 pages. Edited by Kathleen Drzick, Maryellen Hains, and Karla Holloway.

☆Writing Teachers: What We Say About What We Do

Fifteen articles covering a wide range of topics and problems of interest to the secondary and college teacher of writing: how audience affects student writing; looking for the voice in student writing; the in-class journal; teaching writing to minority students; integrating reading and writing; business and science writing; writing across the curriculum; revision; conference and workshop approaches to teaching writing; and varieties of ways to respond and evaluate writing. Concludes with an appendix listing over 50 classroom tested writing assignments. 176 pages. Collected by Stephen Tchudi, edited by Scott McNabb.

☆A Two-Way Street: Reading to Write/Writing to Read

Reviews over eighty books for elementary school children, then lists writing assignments based upon the children's books. Thus, as the authors state, " the regular inclusion of literature in the elementary curriculum can lead to a dynamic interaction and integration in the language arts--writing, reading, speaking, and listening. Developed through a combined effort of the MCTE and the Michigan Department of Education. 105 pages. Edited by Maryellen Hains.

☆Film and Video in the Classroom: A Guide for English Teachers

Eight articles on the issues and practices involved in the teaching of media (films, television), and a guide to classroom use of these media. Also includes sources, resources, and explanations of using these media in the classroom. 63 pages. Edited by Harold Foster.

☆The Language Arts Teacher in Action

Seven articles for the elementary and middle school teacher expressing the importance of: 1) understanding and respecting the child; 2) shaping the curriculum to fit the child; 3) encouraging creativity and divergent thinking; 4) integrating the language arts; and 5) emphasizing language as process. 71 pages. Edited by Constance Weaver and Rollin Douma.

MCTE

P.O. Box 892, Rochester, MI 48063

Books

BEST COPY AVAILABLE